



# *Linked Data Applications:* **There is no One-Size-Fits-All Formula**

**Asunción Gómez-Pérez**

Facultad de Informática, Universidad Politécnica de Madrid  
Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid  
<http://www.oeg-upm.net>  
[asun@fi.upm.es](mailto:asun@fi.upm.es)

## **Acknowledgements:**

O. Corcho, D. Garijo, D. Vila, L. Vilches, B. Villazón

*Work distributed under the license Creative Commons Attribution-Noncommercial-Share Alike 3.0*

1. The concept
2. Foundations
3. The process
4. Examples
  - Libraries: <http://datos.bne.es>
  - Geo: <http://geo.linkeddata.es/>
  - Metereology: <http://aemet.linkeddata.es/>
  - Travelling: <http://webenemasuno.linkeddata.es/>

<http://www.bne.es/>

Cervantes enthusiast from Germany  
visiting Madrid and willing to know  
more about Cervantes' work and life

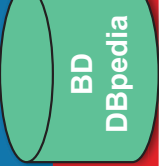
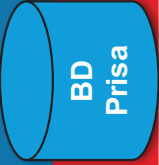
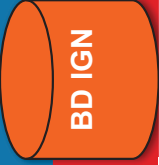
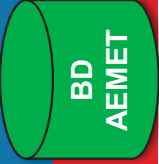
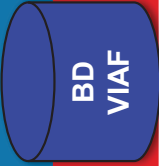
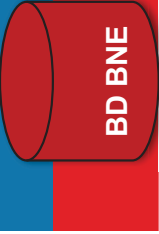
<http://www.viaf.org/>

presas / Ruta de las Tapas

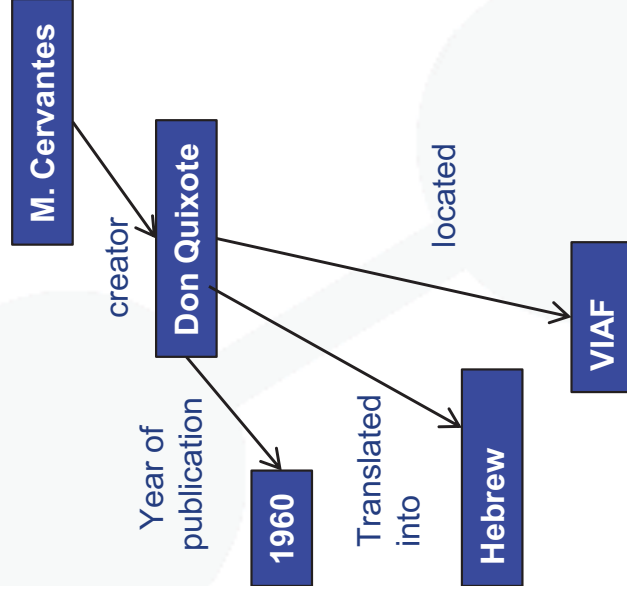
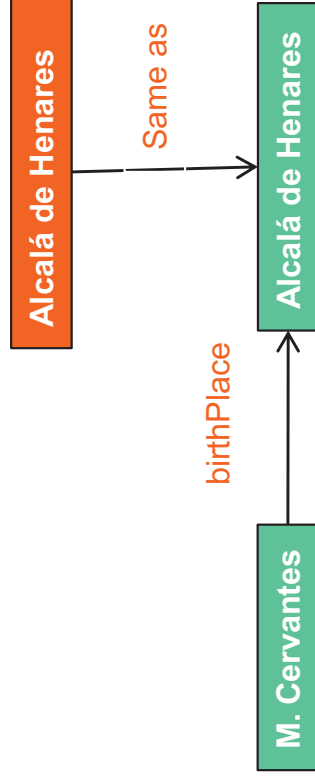
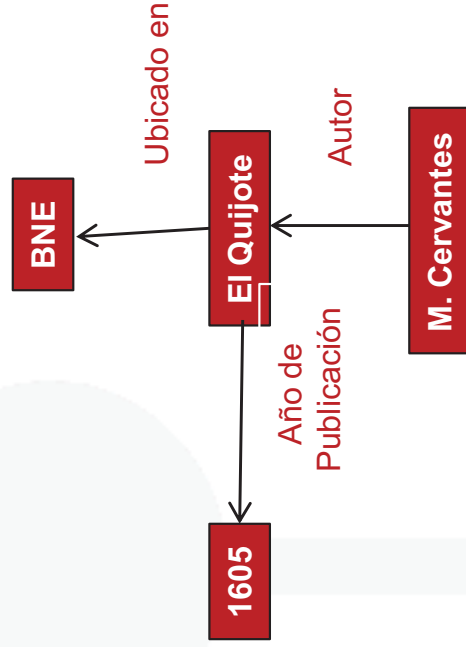
<http://elviajero.elpais.com/>

## Locales

Alcala de Henares. Temperatura (°C)



# Data Integration



1. The concept

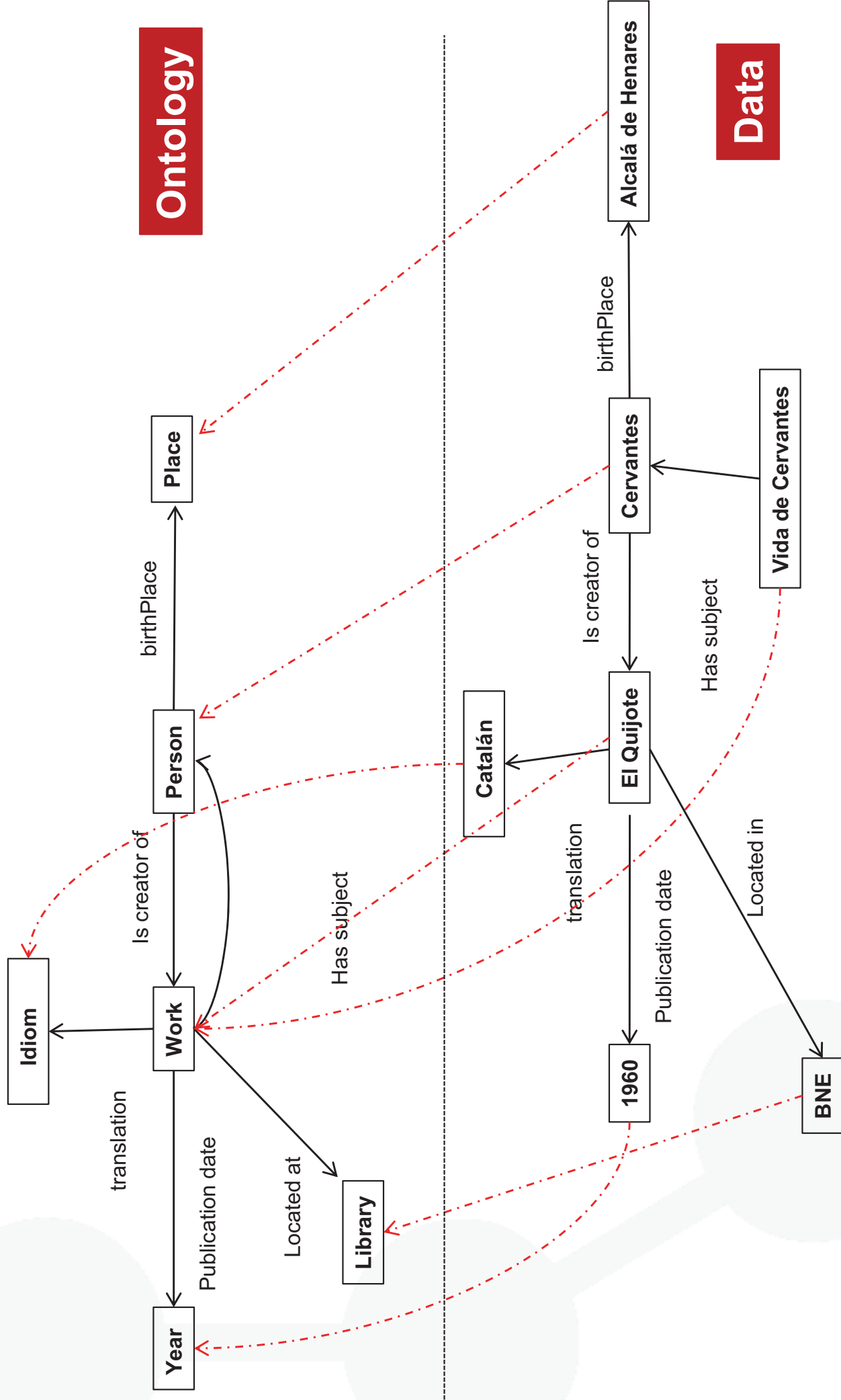
2. Foundations

3. The process

4. Examples

- Libraries: <http://datos.bne.es>
- Geo: <http://geo.linkeddata.es/>
- Metereology: <http://aemet.linkeddata.es/>
- Travelling: <http://webenemasuno.linkeddata.es/>

# The model (Ontology) and the data



# The model (Ontology) and the data

Language

<http://iflstandards.info/ns/fr/frbr/frbrer/C1002>

translation

work

Año

Publication date

<http://iflstandards.info/ns/fr/frbr/frbrer/C1001>

Is creator of

Person

<http://iflstandards.info/ns/fr/frbr/frbrer/C1005>

Located in

Biblioteca

<http://xmlns.com/foaf/0.1/Organization>

Has subject

<http://geo.linkedata.es/ontology/Municipio>

birthPlace

Ontology

Catalán

<http://datos.bne.es/resource/XX1924295>

translation

Don Quijote de la Mancha

1960

Publication date

<http://datos.bne.es/resource/XX3383563>

Es autor

Cervantes Saavedra, Miguel de

<http://datos.bne.es/resource/XX1718747>

Has subject

<http://datos.bne.es/resource/bimo0002045496>

Vida de Miguel de Cervantes Saavedra

Located in

BNE

<http://datos.bne.es/#>

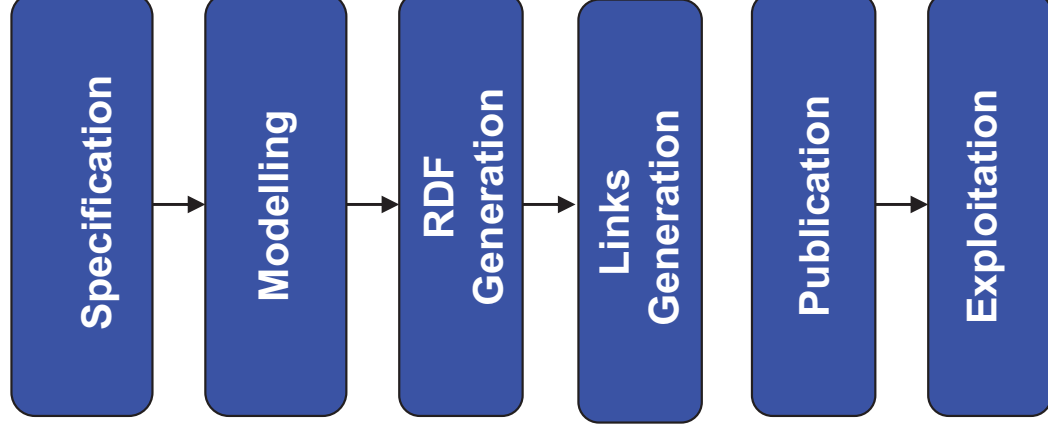
<http://geo.linkedata.es/resource/Alcalá de Henares>

birthPlace

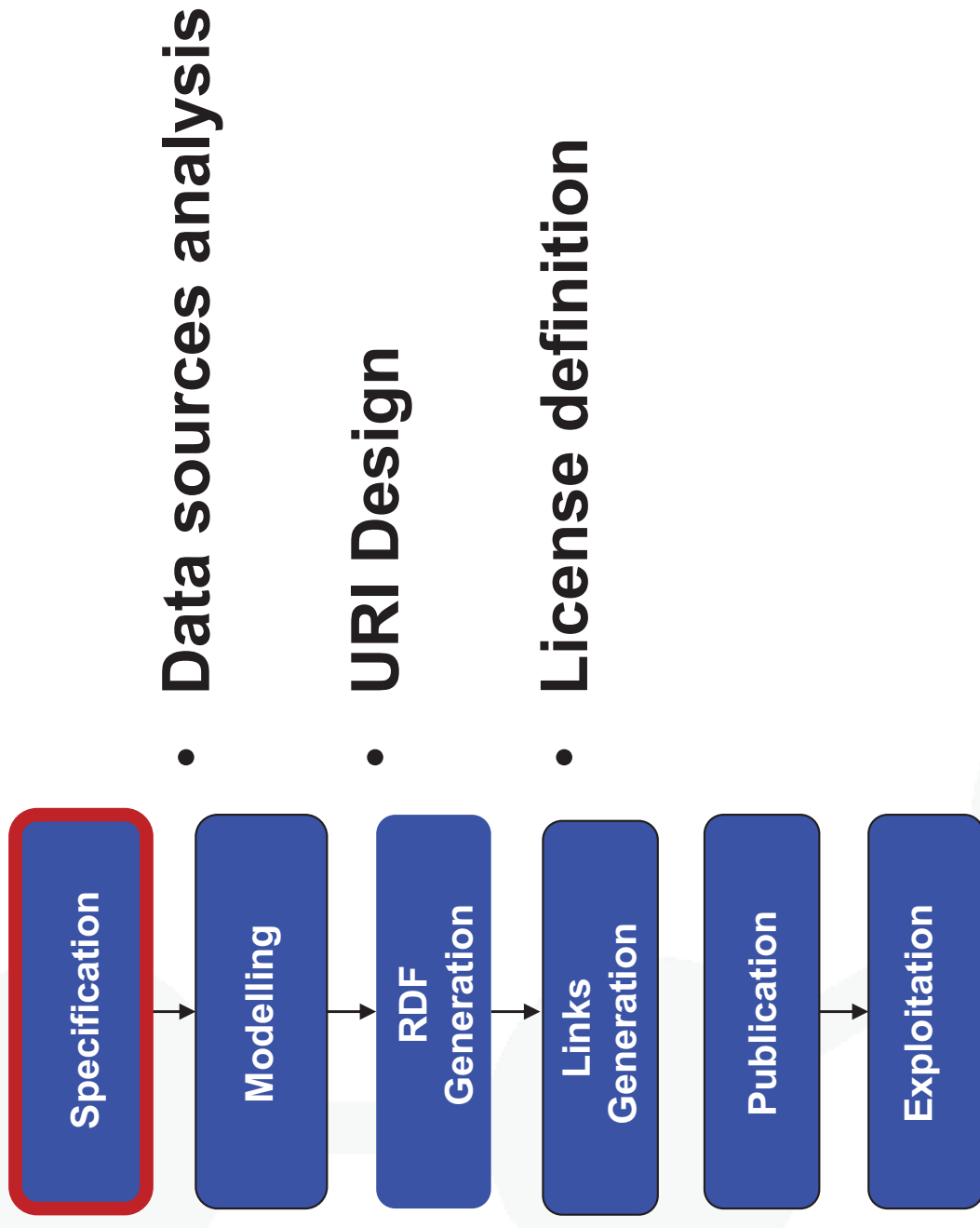
Data

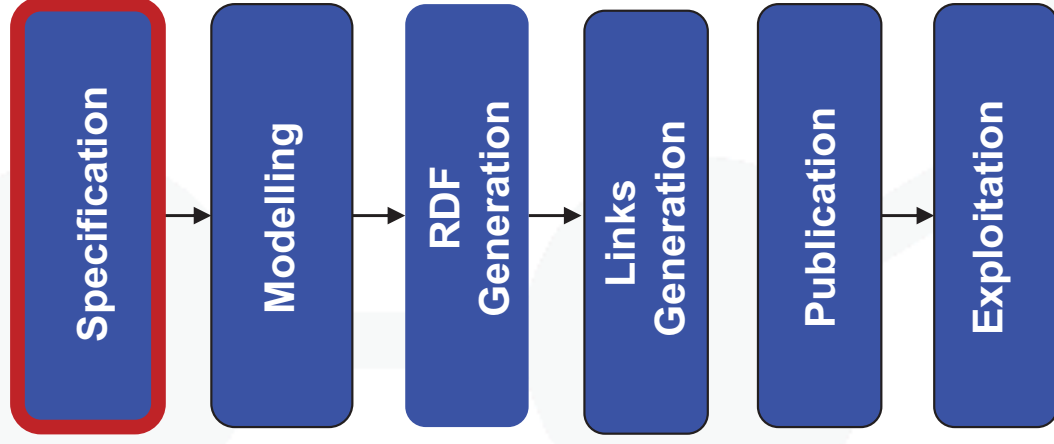


1. The concept
2. Foundations
3. The process
4. Examples

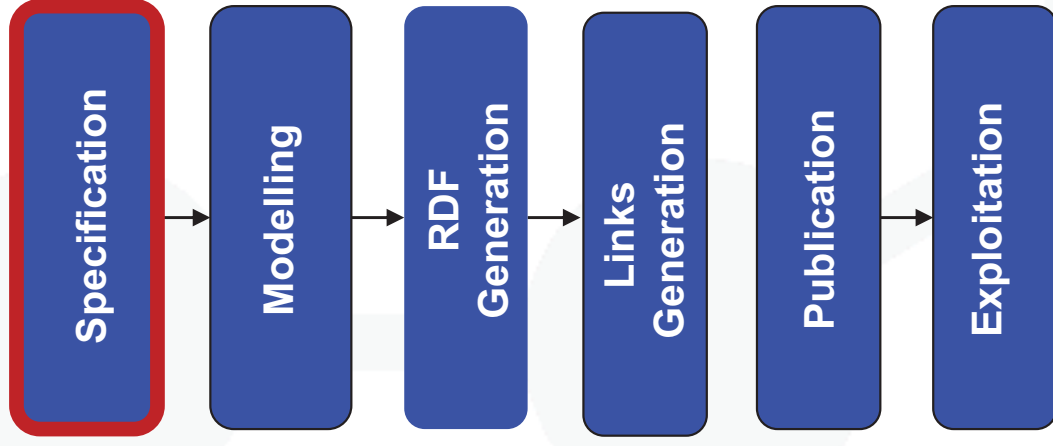








- **Meaningful URIs vs opaque URIs**
- **Separate TBox (ontology model) from ABox**
- **Base URI**
  - <http://linkeddata.es/>
  - <http://geo.linkeddata.es/>
  - <http://otalex.linkeddata.es/>
- **Ontología (TBox URIs)**
  - <http://phenomenontology.linkeddata.es/ontology/{concept|property}>
  - <http://phenomenontology.linkeddata.es/ontology/Municipality>
- **Datos (ABox URIs)**
  - <http://geo.linkeddata.es/resource/{resource type}/{resource name}>
  - <http://geo.linkeddata.es/resource/Municipio/Azuaga>



- **Several possibilities**
  - The UK Open Government License
  - Open Database License
  - Public Domain Dedication and License
  - Open Data Commons Attribution License
  - The Creative Commons Licenses (CC)
- It is also possible to reuse and apply an existing license of the (government) data sources.

**Metadata**

Anon\_0

rdf:type <http://www.w3.org/2004/03/trix/rdfg-1/Graph>

foaf:primaryTopic <http://geo.linkeddata.es/resource/ProvinciaBarcelona>

dc:terms:creator <http://geo.linkeddata.es/resource/organizaci%C3%B3nInstitutoGeogr%C3%A1ficoNacionalDeEspa%C3%B1a>

dc:terms:publisher <http://geo.linkeddata.es>

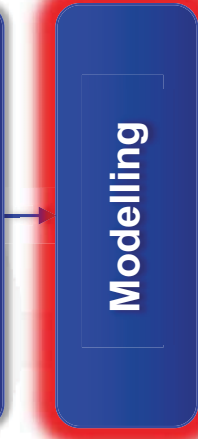
dc:rights La información geográfica digital comprendida en el Equipamiento Geográfico de Referencia Nacional (artículo 1.1 de la Orden FOM/356/2008) así como los Metadatos de los datos geográficos y servicios del IGN-CNIG, no requieren la aceptación de licencia y su uso será, en cualquier caso, libre y gratuito, siempre que se mencione al Instituto Geográfico Nacional como propietario de los datos.

dc:terms:spatial <http://geo.linkeddata.es/resource/Pa%C3%ADsEspana%C3%B1a>

priv:createdBy Anon\_1 (more)

- **Ontologies:**

- A set of terms
- A set of explicit assumptions regarding the intended meaning of the terms.
  - Almost always including concepts and their classification
  - Almost always including properties between concepts



- **Shared understanding** of a domain of interest

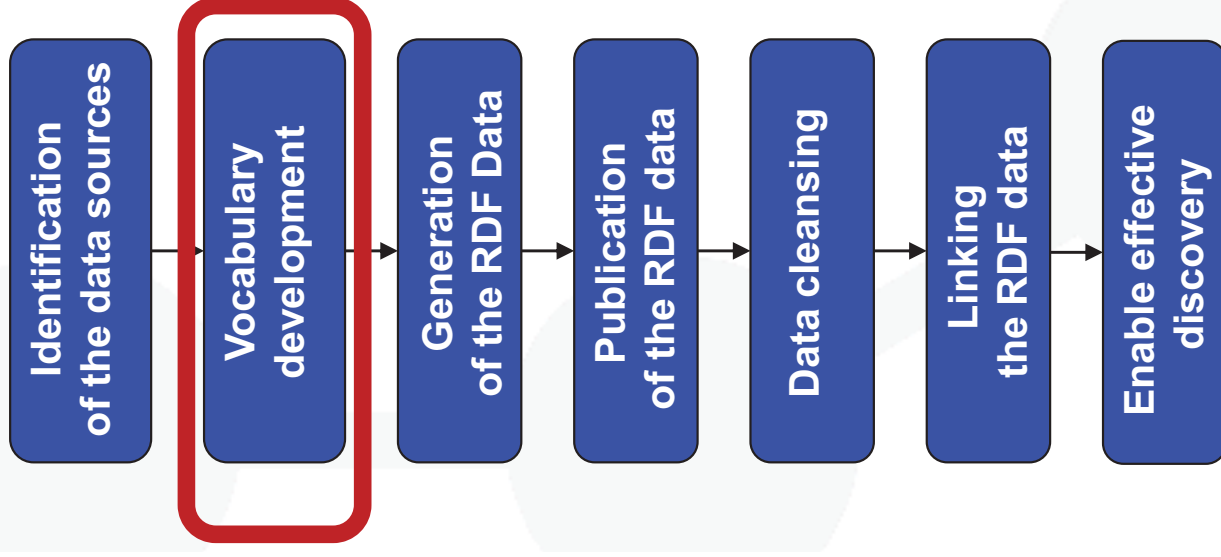


- Ontologies expressed in OWL or **RDF(S)**, both based on RDF

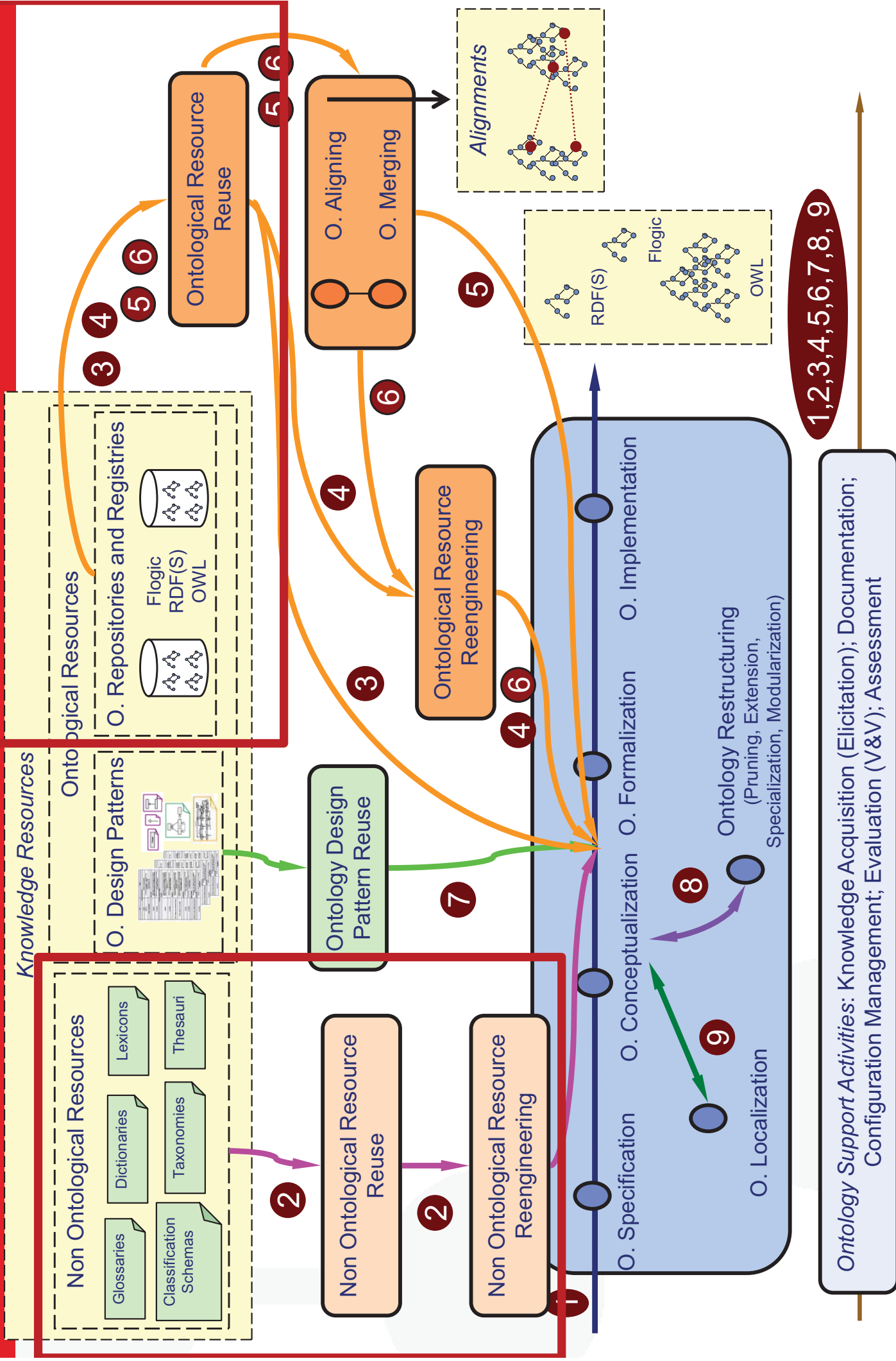


- The **NeOn methodology** helps to build ontologies

## 2. Vocabulary development



- Features
  - Lightweight :
    - Taxonomies and a few properties
  - Consensuated vocabularies
    - To avoid the mapping problems
  - Multilingual
    - Linked data are multilingual
- The NeOn methodology can help to
  - Re-engineer Non ontological resources into ontologies
    - Pros: use domain terminology already consensuated by domain experts
  - Withdraw in heavyweight ontologies those features that you don't need
  - Reuse existing vocabularies



# Reuse available vocabularies



Reuse suitable  
Ontologies and  
vocabularies

...

Search for suitable  
non-ontological resources

are there  
suitable  
resources?

Yes

Build the vocabulary by  
transforming available  
resources

No

Build the vocabulary from  
scratch

Domain-related sites

Highly reliable Web Sites

Government Catalogs

**SchemaWeb**

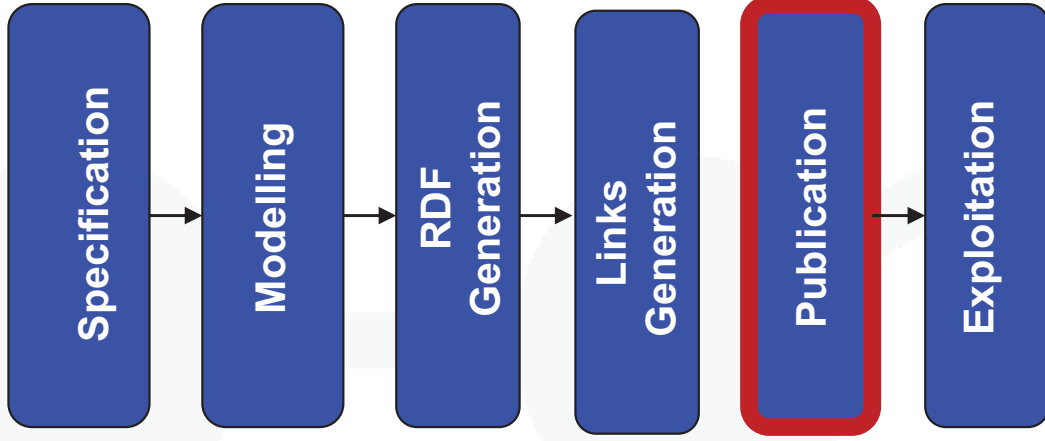


**Swoogle**  
semantic web search  
2007



Linked Open Vocabularies





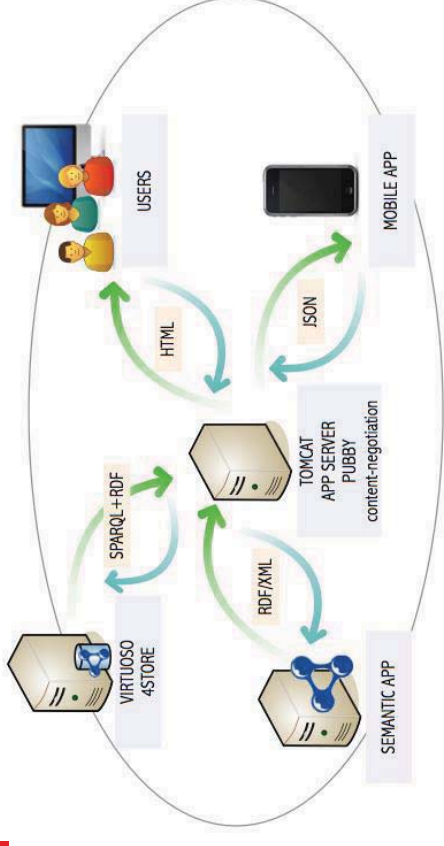
# Publicación

## Data publication

## Metadata publication using VOID

## To facilitate the discovery

- Register in CKAN your dataset
- Use to **sitemap4rdf** to generate the site map
- Upload the site map to **Google** and **Sindice**



## 1. The concept

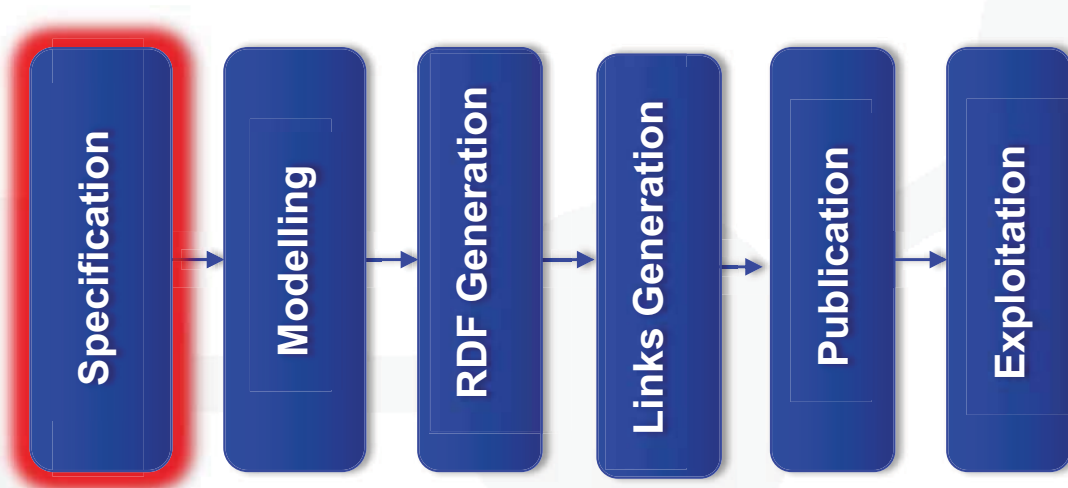
## 2. Foundations

## 3. The process

## 4. Examples

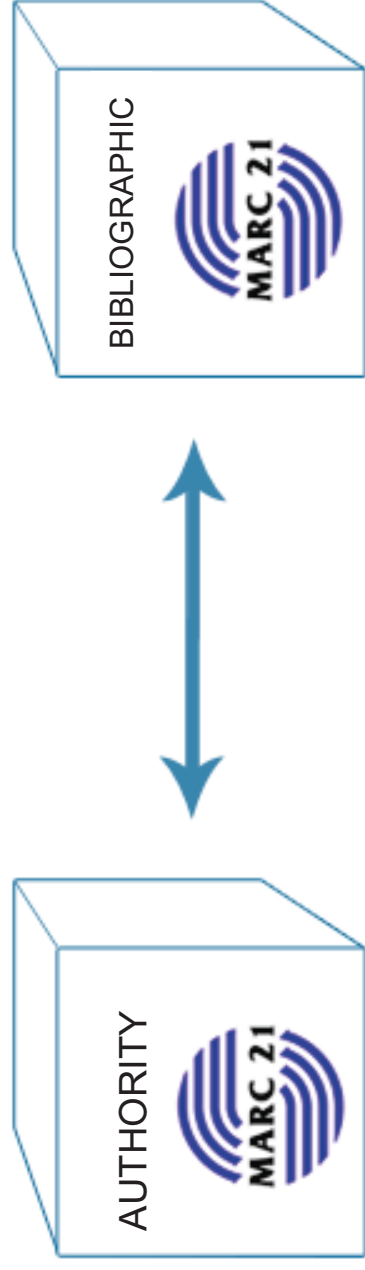
- Libraries: <http://datos.bne.es>
- <http://linkeddata3.dia.fi.upm.es/bne-demo>
- Geo: <http://geo.linkeddata.es/>
- Meteorology: <http://aemet.linkeddata.es/>
- Travelling: <http://webenemasuno.linkeddata.es/>





- Different communication formats:
  - **MARC 21 format for Bibliographic Data**
  - **MARC 21 format for Authority Data**
  - Others: Holdings, Classification, etc.
- Three main elements:
  - **Record structure:** ISO 2709. Fields, indicators, subfields...
  - **Content designation:** "Meaning" of codes and conventions
  - **Content:** Defined outside the MARC standard (ISBD, AACR..)

- Records in the **MARC 21 format**
- **3.9 million** bibliographical records
- **4.2 million** authority records
- Version: November, 2011



2492684	Persons
347737	Corporate bodies
84381	Conferences
715131	Titles
347793	Subject

76576	Maps
320727	Sound recordings
166017	Gravings, drawings, pictures
35770	Manuscripts
143959	Ancient books
2696560	Modern books
178473	Scores
3021	Electronic resources
156634	Serials
96672	Videos

Specification

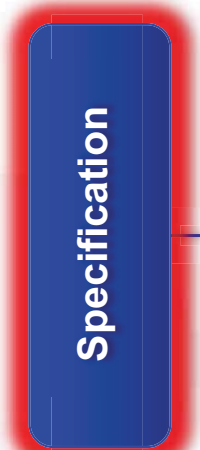
Modelling

RDF Generation

Links Generation

Publication

Exploitation



- Authority record: Camus, Albert\*

Control Field			
001	XX1721208		
005	200012181124		
008	901120nn aijnnaabn n aaa		
016	\$a BNE19900178994		
040	\$a SpMaBN \$b spa \$c SpMaBN \$e rdc \$f embne		
Field	Subfield	Content	HEADING
100	10 \$a	Camus, Albert	1XX
	\$d	1913-1960	
670 \$a El mite de Sísif, 1987 \$b port. (Albert Camus)			
670 \$a Dic. de filosofía, de J. Ferrater Mora, 1980\$b(Camus., Albert (1913-1960); n. Mondovi, Argel)			
670 \$a Aut. BN-OPALE, 1995 \$b (Camus, Albert)			

\* <https://datos.bne.es/catalogo/XX1721208>

\* <http://datos.bne.es/resource/XX1721208>

# MARC21 record content designation

- Authority record: Camus, Albert\*

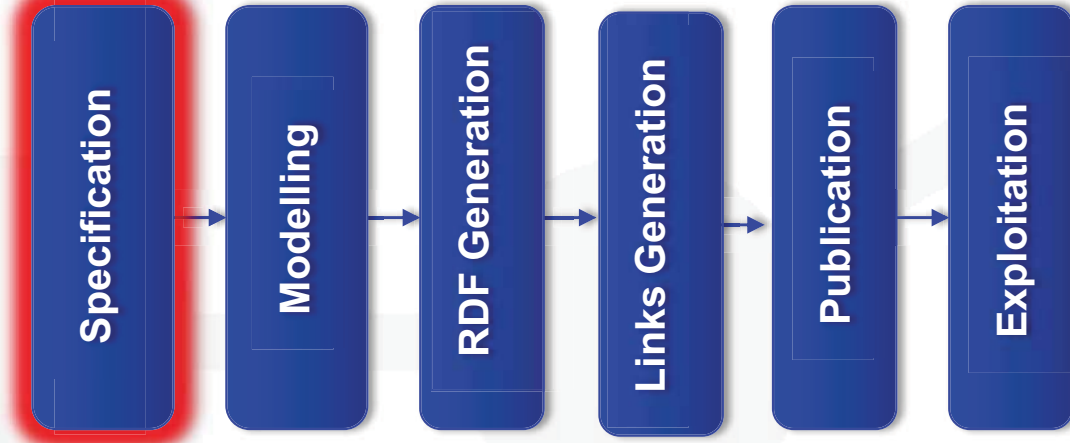
Control Number	001	XX1721208	HEADING – Personal	
Personal name	100	10 \$a	Camus, Albert	Name 100
Dates associated with name		\$d	1913-1960	
Source consulted	670	\$a	El mite de Sísif, 1987 \$b port. (Albert Camus)	
Citation				

- Human reading:

*An authority record that describes a **Person**, named **Camus, Albert** with associated dates **1913-1960***

\* <http://datos.bne.es/resource/XX1721208>

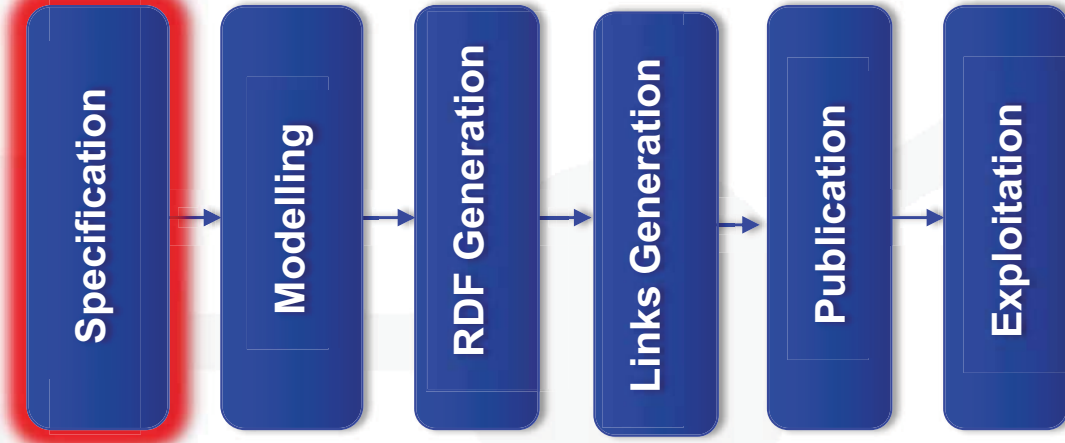
# Frequency of codes in records



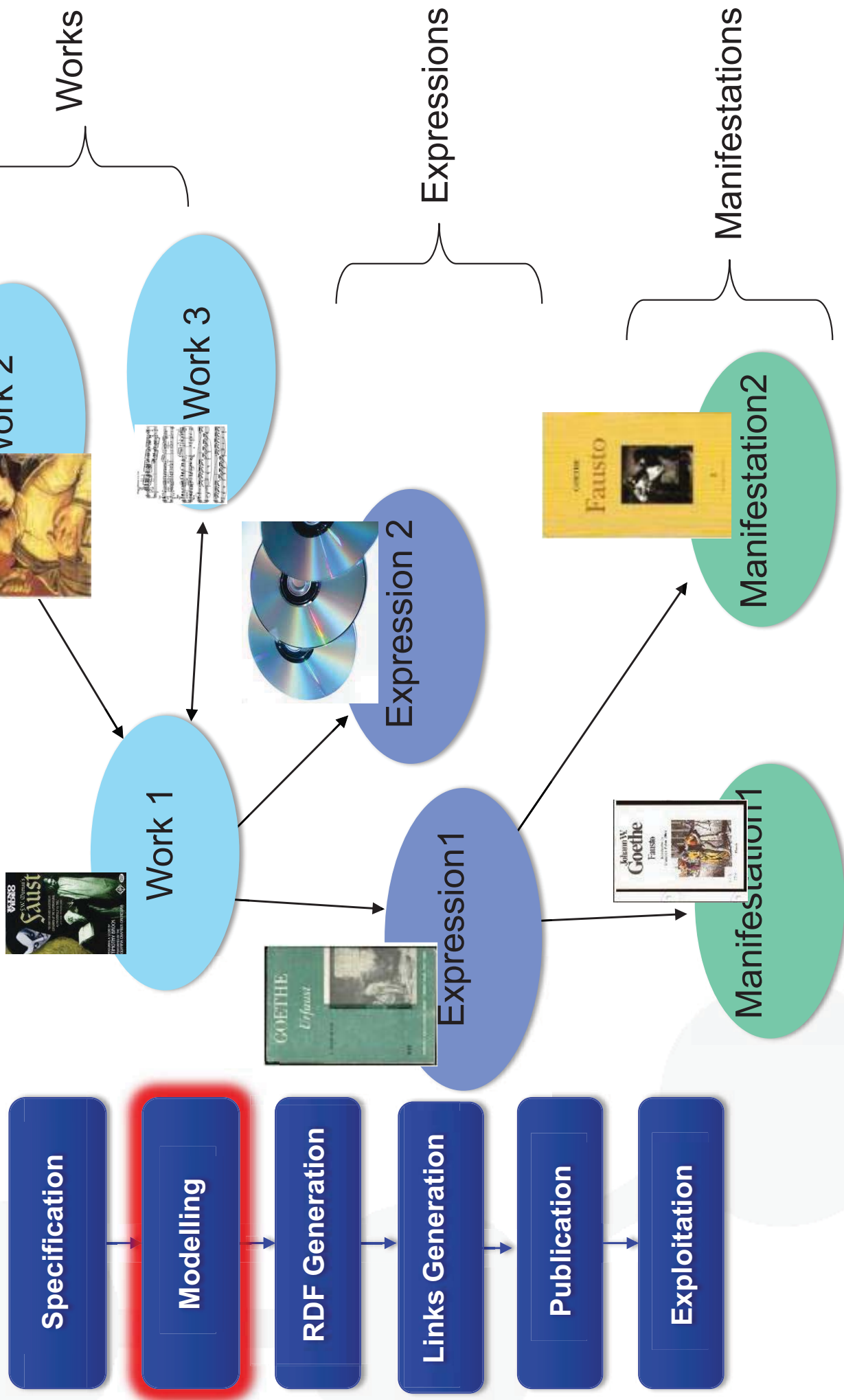
	A	B	C
1	Combination	Records count	Example
2	a	881115	[\$aYcardo, Maruja]
3	at	685861	[\$aGual, Adriat, \$t Donzell qui cerca muller]
4	adt	347783	[\$aVargas, Pedro, \$d1906-1989, \$tCanciones de nuestra vida]
5	ad	183377	[\$aComette de Saint Cyr Pierre, \$d1939.1]
6	atl	87828	
7	adtl	60231	
8	ae	35936	
9	ac	25992	
10	adv	19807	
11	act	17359	
12	ade	14573	
13	av	11427	
14	aq	10143	
15	adx	8326	
16	adtp	8167	
17	ax	6563	
18	acd	6440	
19	acdt	6262	
20	aqt	6133	
21	adto	3480	
22	adtmnr	3340	
23	adtk	3085	
24	adtpo	2964	
25	ace	2479	
26	adtrn	2360	
27	adtn	2330	
28	adtnp	2220	
29	adxv	2155	
30	adtm	1990	
31	adtmr	150	
32	adtmr	151	
33	adtmr	110	
34	adtmr	111	
35	adtmr	112	
36	adtmr	113	
37	adtmr	114	
38	adtmr	115	
39	adtmr	116	
40	adtmr	117	
41	adtmr	118	
42	adtmr	119	
43	adtmr	120	
44	adtmr	121	
45	adtmr	122	
46	adtmr	123	
47	adtmr	124	
48	adtmr	125	
49	adtmr	126	
50	adtmr	127	
51	adtmr	128	
52	adtmr	129	
53	adtmr	130	
54	adtmr	131	
55	adtmr	132	
56	adtmr	133	
57	adtmr	134	
58	adtmr	135	
59	adtmr	136	
60	adtmr	137	
61	adtmr	138	
62	adtmr	139	
63	adtmr	140	
64	adtmr	141	
65	adtmr	142	
66	adtmr	143	
67	adtmr	144	
68	adtmr	145	
69	adtmr	146	
70	adtmr	147	
71	adtmr	148	
72	adtmr	149	
73	adtmr	150	
74	adtmr	151	
75	adtmr	152	
76	adtmr	153	
77	adtmr	154	
78	adtmr	155	
79	adtmr	156	
80	adtmr	157	
81	adtmr	158	
82	adtmr	159	
83	adtmr	160	
84	adtmr	161	
85	adtmr	162	
86	adtmr	163	
87	adtmr	164	
88	adtmr	165	
89	adtmr	166	
90	adtmr	167	
91	adtmr	168	
92	adtmr	169	
93	adtmr	170	
94	adtmr	171	
95	adtmr	172	
96	adtmr	173	
97	adtmr	174	
98	adtmr	175	
99	adtmr	176	
100	adtmr	177	
101	adtmr	178	
102	adtmr	179	
103	adtmr	180	
104	adtmr	181	
105	adtmr	182	
106	adtmr	183	
107	adtmr	184	
108	adtmr	185	
109	adtmr	186	
110	adtmr	187	
111	adtmr	188	
112	adtmr	189	
113	adtmr	190	
114	adtmr	191	
115	adtmr	192	
116	adtmr	193	
117	adtmr	194	
118	adtmr	195	
119	adtmr	196	
120	adtmr	197	
121	adtmr	198	
122	adtmr	199	
123	adtmr	200	
124	adtmr	201	
125	adtmr	202	
126	adtmr	203	
127	adtmr	204	
128	adtmr	205	
129	adtmr	206	
130	adtmr	207	
131	adtmr	208	
132	adtmr	209	
133	adtmr	210	
134	adtmr	211	
135	adtmr	212	
136	adtmr	213	
137	adtmr	214	
138	adtmr	215	
139	adtmr	216	
140	adtmr	217	
141	adtmr	218	
142	adtmr	219	
143	adtmr	220	
144	adtmr	221	
145	adtmr	222	
146	adtmr	223	
147	adtmr	224	
148	adtmr	225	
149	adtmr	226	
150	adtmr	227	
151	adtmr	228	
152	adtmr	229	
153	adtmr	230	
154	adtmr	231	
155	adtmr	232	
156	adtmr	233	
157	adtmr	234	
158	adtmr	235	
159	adtmr	236	
160	adtmr	237	
161	adtmr	238	
162	adtmr	239	
163	adtmr	240	
164	adtmr	241	
165	adtmr	242	
166	adtmr	243	
167	adtmr	244	
168	adtmr	245	
169	adtmr	246	
170	adtmr	247	
171	adtmr	248	
172	adtmr	249	
173	adtmr	250	
174	adtmr	251	
175	adtmr	252	
176	adtmr	253	
177	adtmr	254	
178	adtmr	255	
179	adtmr	256	
180	adtmr	257	
181	adtmr	258	
182	adtmr	259	
183	adtmr	260	
184	adtmr	261	
185	adtmr	262	
186	adtmr	263	
187	adtmr	264	
188	adtmr	265	
189	adtmr	266	
190	adtmr	267	
191	adtmr	268	
192	adtmr	269	
193	adtmr	270	
194	adtmr	271	
195	adtmr	272	
196	adtmr	273	
197	adtmr	274	
198	adtmr	275	
199	adtmr	276	
200	adtmr	277	
201	adtmr	278	
202	adtmr	279	
203	adtmr	280	
204	adtmr	281	
205	adtmr	282	
206	adtmr	283	
207	adtmr	284	
208	adtmr	285	
209	adtmr	286	
210	adtmr	287	
211	adtmr	288	
212	adtmr	289	
213	adtmr	290	
214	adtmr	291	
215	adtmr	292	
216	adtmr	293	
217	adtmr	294	
218	adtmr	295	
219	adtmr	296	
220	adtmr	297	
221	adtmr	298	
222	adtmr	299	
223	adtmr	300	
224	adtmr	301	
225	adtmr	302	
226	adtmr	303	
227	adtmr	304	
228	adtmr	305	
229	adtmr	306	
230	adtmr	307	
231	adtmr	308	
232	adtmr	309	
233	adtmr	310	
234	adtmr	311	
235	adtmr	312	
236	adtmr	313	
237	adtmr	314	
238	adtmr	315	
239	adtmr	316	
240	adtmr	317	
241	adtmr	318	
242	adtmr	319	
243	adtmr	320	
244	adtmr	321	
245	adtmr	322	
246	adtmr	323	
247	adtmr	324	
248	adtmr	325	
249	adtmr	326	
250	adtmr	327	
251	adtmr	328	
252	adtmr	329	
253	adtmr	330	
254	adtmr	331	
255	adtmr	332	
256	adtmr	333	
257	adtmr	334	
258	adtmr	335	
259	adtmr	336	
260	adtmr	337	
261	adtmr	338	
262	adtmr	339	
263	adtmr	340	
264	adtmr	341	
265	adtmr	342	
266	adtmr	343	
267	adtmr	344	
268	adtmr	345	
269	adtmr	346	
270	adtmr	347	
271	adtmr	348	
272	adtmr	349	
273	adtmr	350	
274	adtmr	351	
275	adtmr	352	
276	adtmr	353	
277	adtmr	354	
278	adtmr	355	
279	adtmr	356	
280	adtmr	357	
281	adtmr	358	
282	adtmr	359	
283	adtmr	360	
284	adtmr	361	
285	adtmr	362	
286	adtmr	363	
287	adtmr	364	
288	adtmr	365	
289	adtmr	366	
290	adtmr	367	
291	adtmr	368	
292	adtmr	369	
293	adtmr	370	
294	adtmr	371	
295	adtmr	372	
296	adtmr	373	
297	adtmr	374	
298	adtmr	375	
299	adtmr	376	
300	adtmr	377	
301	adtmr	378	
302	adtmr	379	
303	adtmr	380	
304	adtmr	381	
305	adtmr	382	
306	adtmr	383	
307	adtmr	384	
308	adtmr	385	
309	adtmr	386	
310	adtmr	387	
311	adtmr	388	
312	adtmr	389	
313	adtmr	390	
314	adtmr	391	
315	adtmr	392	
316	adtmr	393	
317	adtmr	394	
318	adtmr	395	
319	adtmr	396	
320	adtmr	397	
321	adtmr	398	
322	adtmr	399	
323	adtmr	400	
324	adtmr	401	
325	adtmr	402	
326	adtmr	403	
327	adtmr	404	
328	adtmr	405	
329	adtmr	406	
330	adtmr	407	
331	adtmr	408	
332	adtmr	409	
333	adtmr	410	
334	adtmr	411	
335	adtmr	412	
336	adtmr	413	
337	adtmr	414	
338	adtmr	415	
339	adtmr	416	
340	adtmr	417	
341	adtmr	418	
342	adtmr	419	
343	adtmr	420	
344	adtmr	421	
345	adtmr	422	
346	adtmr	423	
347	adtmr	424	
348	adtmr	425	
349	adtmr	426	
350	adtmr	427	
351	adtmr	428	
352	adtmr	429	
353	adtmr	430	
354	adtmr	431	
355	adtmr	432	
356	adtmr	433	
357	adtmr	434	
358	adtmr	435	
359	adtmr	436	
360	adtmr	437	
361	adtmr	438	
362	adtmr	439	
363	adtmr	440	
364	adtmr	441	
365	adtmr	442	
366	adtmr	443	
367	adtmr	444	
368	adtmr	445	
369	adtmr	446	
370	adtmr	447	
371	adtmr	448	
372	adtmr	449	
373	adtmr	450	
374	adtmr	451	
375	adtmr	452	
376	adtmr	453	
377	adtmr	454	
378	adtmr	455	
379	adtmr	456	
380	adtmr	457	
381	adtmr	458	
382	adtmr	459	
383	adtmr	460	
384	adtmr	461	
385	adtmr	462	
386	adtmr	463	
387	adtmr	464	
388</			



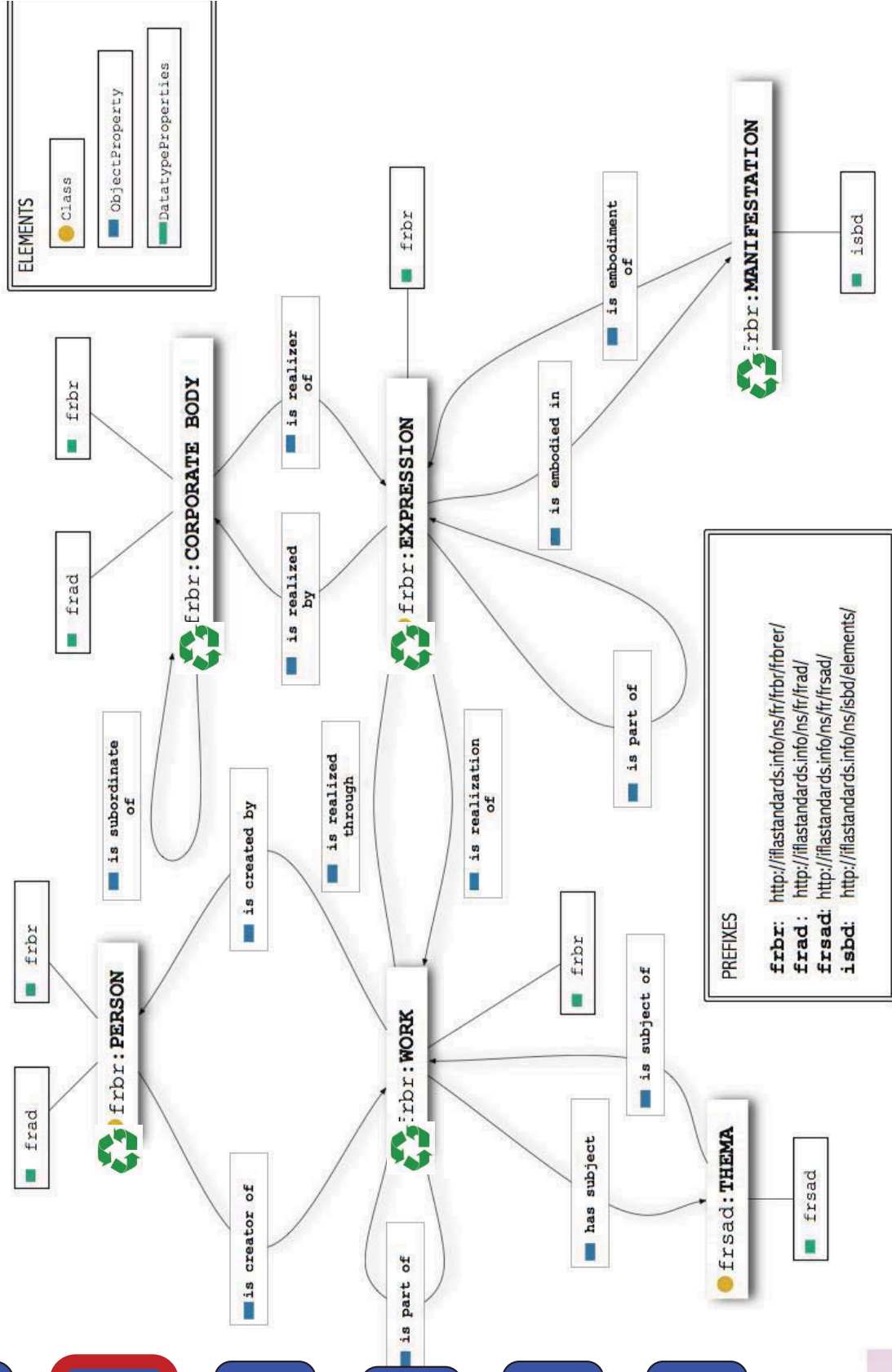
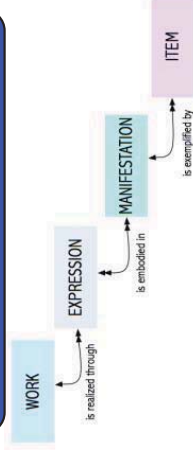
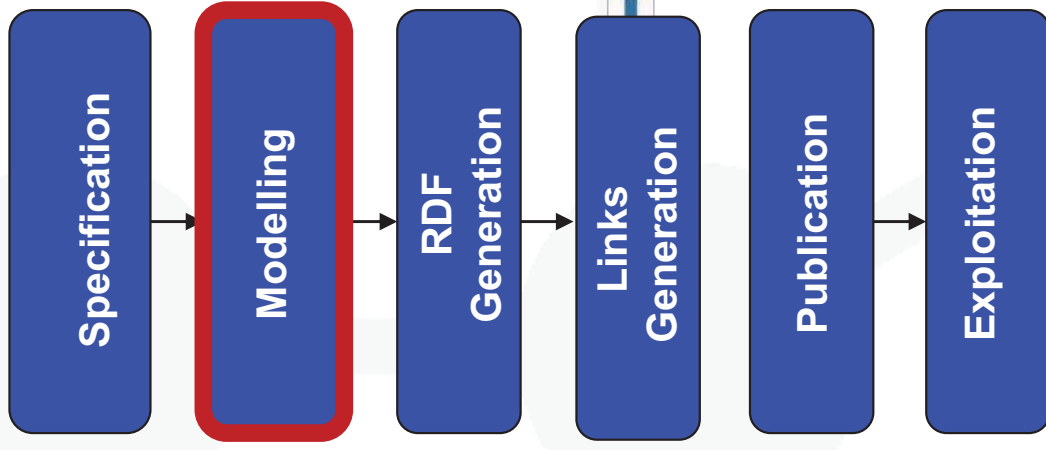
- **Source data:** MARC 21 records, not RDB. Very flat structure difficult to map to richer models
- Domain experts (catalogers) need to be part of the mapping process.
- Data quality good but still many errors: **reporting**.
- Iterative and incremental transformation process: **measure coverage and progress**.
- Highly specialized library models: FRBR, ISBD.
- Multilinguality, collaboration with IFLA




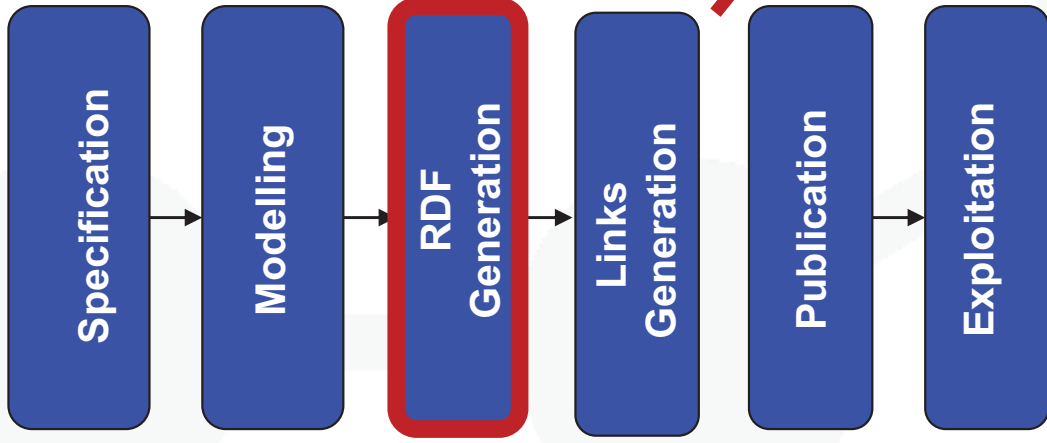
# Model: FRBR at a glance



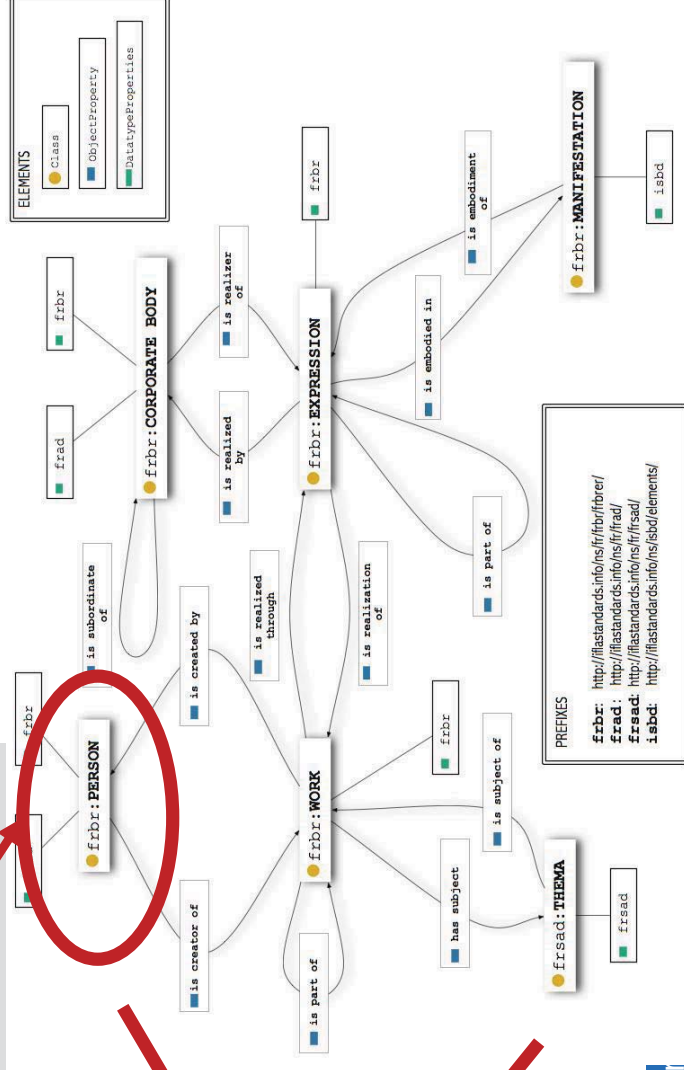
# The Ontology: based on IFLA vocabularies



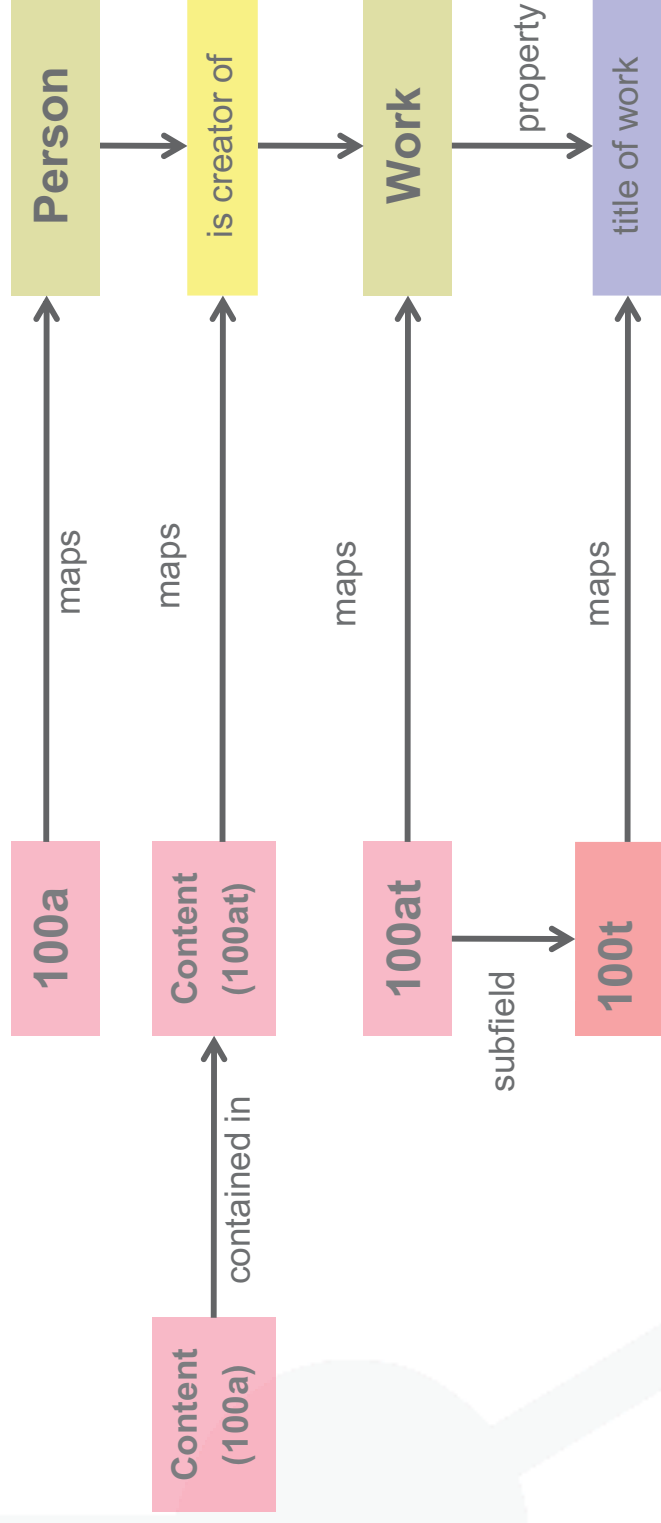
# Who will be the mapping generator?



001 XX1721208  
005 200012181124  
008 901120nn aijnnaabn n aaa  
016 \$a BNE19900178994  
040 \$a SpMaBN \$b spa \$c SpMaBN \$e rdc \$f embne  
100 10 \$a Camus, Albert  
\$d 1913-1960  
670 \$a El mite de Sísif, 1987 \$b port. (Albert Camus)  
670 \$a Dic. de filosofía, de J. Ferrater Mora,  
1980\$b(Camus., Albert (1913-1960); n. Mondovi,  
Argel)  
670 \$a Aut. BN-OPALE, 1995 \$b (Camus, Albert)

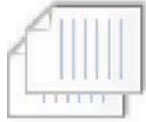


# Similar to mapping ontologies

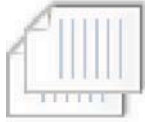




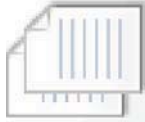
- Three spreadsheets:



Classification  
mapping



Annotation  
mapping



Relationships  
mapping



## Basic structure

MARC21 info	Records count	Content sample	Mapping
100 \$a \$d	888.880	Camus, Albert 1913-1960	foaf:Person
100 \$a	999.999	Cervantes, Miguel de	foaf:name
100 \$a \$m	10.000	Cervantes, iguel	<b>ERROR</b>

# Librarians create mappings using excell

## Classification mapping



	A	B	C	
1	Combination	Records count	Example	Mapping value
2	a	881115	[\$aYcardo, Maruja]	Person
3	at	685861	[\$aGual, Adrià, \$t Donzell qui cerca muller]	Work
4	adt	347783	[\$aVargas, Pedro, \$d1906-1989, \$tCanciones de nuestra vida]	Work
5	ad	183377	[\$aCornette de Saint Cyr, Pierre, \$d1939-]	Person
6	atl	87828	[\$aJames, Samantha, \$tSterling family trilogy, \$IEspaOl (Barcelona)]	Expression
7	adtl	60231	[\$aSalvatore, R. A., \$d1959-, \$tThe sellwords, \$IEspaOl (Barcelona)]	Expression
8	ae	35936	[\$aAbad, Atanasio, \$eimp.]	Person
9	ac	25992	[\$aA. M. B., \$cdottor]	Person
10	adv	19807	[\$aCavanna, Xavier, \$d1961-, \$vExposiciones]	Thema
11	act	17359	[\$aCoelho, J. R., \$c(m'ico), \$tBasta amor meu terno peito]	Work
12	ade	14573	[\$aMontoliu, Tete, \$d1933-1997, \$eint.]	Person
13	av	11427	[\$aBoisvert, Georges, \$vHomenajes]	Thema
14	aq	10143	[\$aValdÈs, Manuel, \$qValdÈs García]	Person
15	adx	8326	[\$aFunes, Gregorio, \$d1749-1829, \$xBiografías]	Thema
16	adtp	8167	[\$aBenevoli, Orazio, \$d1605-1672, \$tDeus pro nobis, \$pSanctus]	Work
17	ax	6563	[\$aGarret, Almeida, \$xCrítica e interpretação]	Thema
18	acd	6440	[\$aAbad, Bautista, \$c(O.S.A.), \$d -ca. 1701]	Person
19	acdt	6262	[\$aTom's Moro., \$cSanto, \$d1478-1535, \$tThe last things]	Work
20	aqt	6133	[\$aIvarez, Pilar, \$qIvarez Saenz, \$tConvivir]	Work
21	adto	3480	[\$aTorres Nin, Lorenzo, \$d1890-1964, \$tLas pobres millonarias, \$oarr.]	Expression
22	adtmnr	3340	[\$aLeybach, Ignace Xavier Joseph, \$d1817-1891, \$tNocturnos., \$mpiano, \$r	Work
23	adtk	3085	[\$aChaikovski, Piotr Ilich, \$d1840-1893, \$tEl lago de los cisnes (suite), \$k	Expression
24	adtpto	2964	[\$aDauphin, Léopold, \$d1847-1925, \$tDjanika, \$pNoble Époux, \$oarr.]	Expression
25	ace	2479	[\$aAcosta, Felipe de, \$c(O.P.), \$eed. lit]	Person
26	adtmn	2360	[\$aLigeti, György, \$d1923-2006, \$tCuartetos, \$m cuerda, \$nn. 1]	Work
27	adtn	2330	[\$aBernaola, Carmelo, \$d1929-2002, \$tSinfonías, \$nn. 3]	Work
28	adtnp	2220	[\$aBach, Johann Sebastian, \$d1685-1750, \$tMatthäuspassion., \$nN. 20.,	Work
29	adxv	2155	[\$aFiguera Aymerich, Ingela, \$d1902-1984, \$xHomenajes, \$vIdeograbacio	Thema
30	adtm	1990	[\$aGorriti, Felipe, \$d1839-1896, \$tMagnificat. \$m4 voces. orquestal	Work



# Librarians create mappings using excell

## Annotation

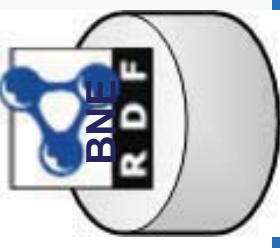
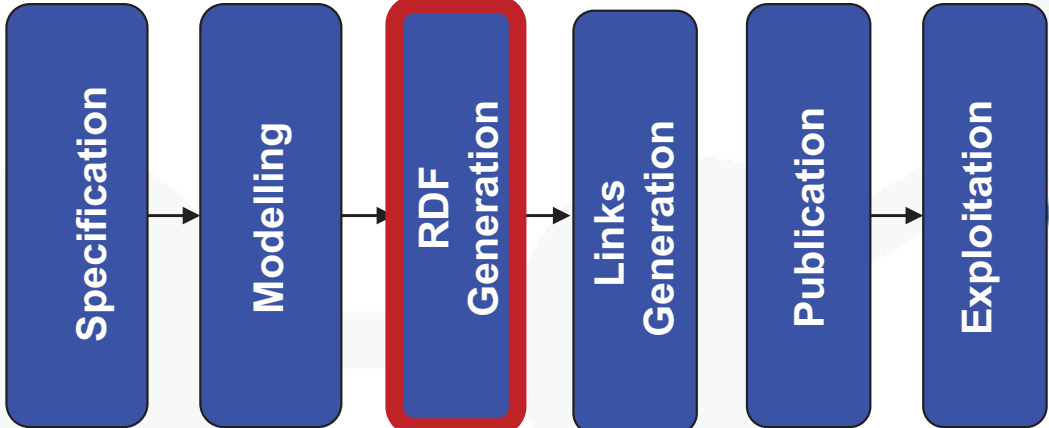
		A	B	C	D
1		Combinatio	Records c	Example	Mapping value
2	260b		3892943	Imp. Lemerrier et C.ie	<a href="http://iflastandards.info/ns/isbd/elements/P1117">http://iflastandards.info/ns/isbd/elements/P1117</a>
3	245a		3871272	Polka elegante para piano	<a href="http://iflastandards.info/ns/isbd/elements/P1004">http://iflastandards.info/ns/isbd/elements/P1004</a>
4	040a		3870286	SpMaBN	
5	260a		3845589	[Paris	<a href="http://iflastandards.info/ns/isbd/elements/P1022">http://iflastandards.info/ns/isbd/elements/P1022</a>
6	300a		3827202	2 p.	<a href="http://iflastandards.info/ns/isbd/elements/P1185">http://iflastandards.info/ns/isbd/elements/P1185</a>
7	245h		3751432	M' sica impresa	<a href="http://iflastandards.info/ns/isbd/elements/P1018">http://iflastandards.info/ns/isbd/elements/P1018</a>
8	260c		3714336	1865-1869]	<a href="http://iflastandards.info/ns/isbd/elements/P1018">http://iflastandards.info/ns/isbd/elements/P1018</a>
9	300c		3526395	22 cm	<a href="http://iflastandards.info/ns/isbd/elements/P1018">http://iflastandards.info/ns/isbd/elements/P1018</a>
10	016a		3103593	bipaBNE20070006028	
11	245c		2805186	por Henri Cramer	
12	017a		2589626	GC 30-2006	
13	017b		2582999	Oficina Depósito Legal MΣlaga	
14	6502		1707468	embne	
15	300b		1627566	aguafuerte	
16	020a		1527638	84-7568-355-X	
17	016a		1461610	ININQ0000366704E	

		B		C	D
1		Variation	Sample	Relationship	
2	n		Sociedad Didáctico-Musical Método especial de solfeo cor		Is part of work
3	o		Bach, Johann Sebastian 1685-1750 Partita laúd		
4	p		Iglesia Católica Synodus Episcoporum (2008) Verbum Dor		frbr:P2057; frbr:P2058
5	q		Isabel María Isabel María López Campos The face		
6	r		Chopin, Fryderyk 1810-1849 Valses piano mi menor		
7	s		Beethoven, Ludwig van 1770-1827 Fantasien pia		
8	t		Castellà Andreu, Josep M. La función constitui		

Relationships  
mapping



# Marimba interprets the Mappings and generate the RDF



```
001 XX1721208
.....
100 10 $a Camus, Albert
      $d 1913-1960
.....
```



- **Classify:** Exploiting the heading field and subfield codes.  
100 \$a \$d → Person (it has a personal name)  
100 \$a \$d \$t → Work (it has a title)
- **Annotate:** Using subfield codes and the **content**.  
100 \$a "Camus, Albert" → frbr:3001 "Camus, Albert"  
100 \$t "La Peste" → frbr:P3039 "La Peste"

MARC 21 record (Input)	Action	RDF (Output)
100 \$a \$d	Classify	rdf:type frbr:C1005
100 \$a Camus, Albert	Annotate	frbr:P3039 "Camus, Albert"
100 \$d 1913-1960	Annotate	frbr:P3040 "1913-1960"

# Mapping process more in detail

- **But**, what about the **relationships** between the entities?
- Relationships between records are **not explicit** in MARC.

Goal: The work "La Peste" was created by Albert Camus

# RDF Generation

001 XX1721208

100 10 \$a Camus, Albert \$d 1913-1960

# Common

001 XX1910518

100 10 \$a Camus, Albert\$d1913-1960

# Common

## \$t la peste

# Diff

# Work

We know the type of R1 and R2, and we look at the heading diff

bne:XX1721208      **frbr:2010**      bne:XX1910518  
(isCreatorOf)

\* <http://datos.bne.es/resource/XX1910518>

# Marimba: Mapping process summary

Specification

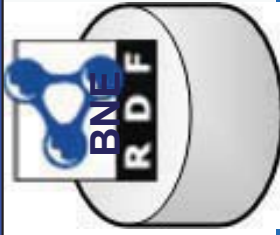
Modelling

RDF  
Generation

Links  
Generation

Publication

Exploitation



(MARC records)

```
001 XX1721208
100 10 $a Camus, Albert $d 1913-1960
```

Classify

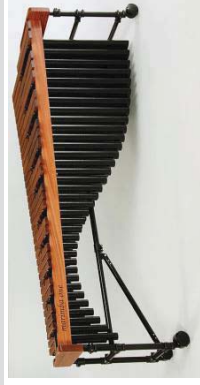
```
bne:XX1721208 a frbr:Person
```

Annotate

```
bne:XX1721208 a frbr:Person
frbr:name "Camus, Albert" .
frbr:hasDates 1913-1960
```

Relate

```
bne:XX1721208 a frbr:Person
frbr:name "Camus, Albert" .
frbr:hasDates 1913-1960 .
frbr:isCreatorOf bne:XX1721208
```



```
001 XX1910518
100 10 $a Camus, Albert$d1913-1960 $tLa
    peste
```

```
bne:XX1910518 a frbr:Work
```

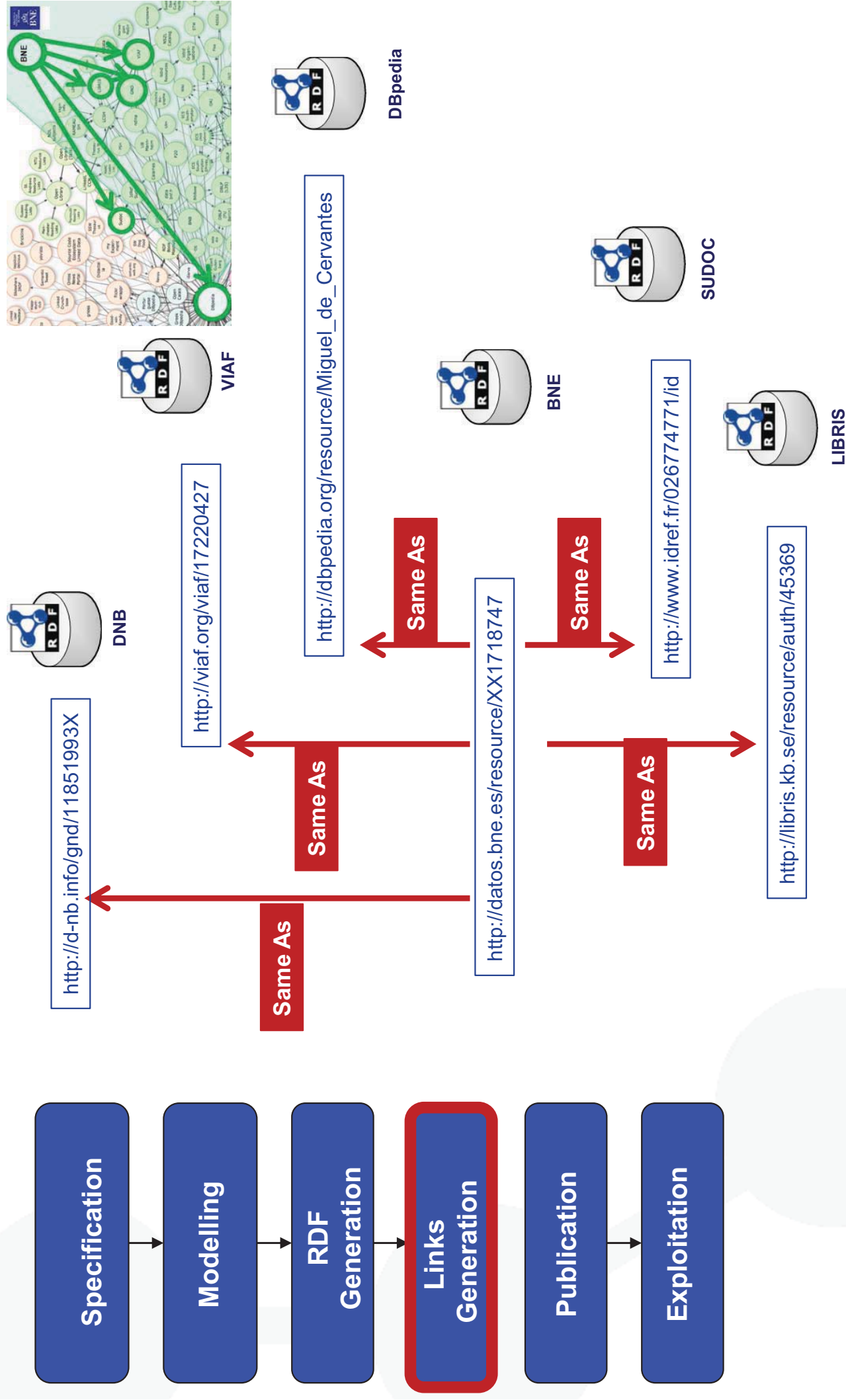
```
bne:XX1910518 a frbr:Work
frbr:title "La Peste"
```

```
bne:XX1910518 a frbr:Work
    frbr:title "La Peste" .
frbr:isCreatedBy bne:XX1721208
```





# Marimba links with other resources: VIAF, DNB, SUDOC, LIBRIS, DBpedia



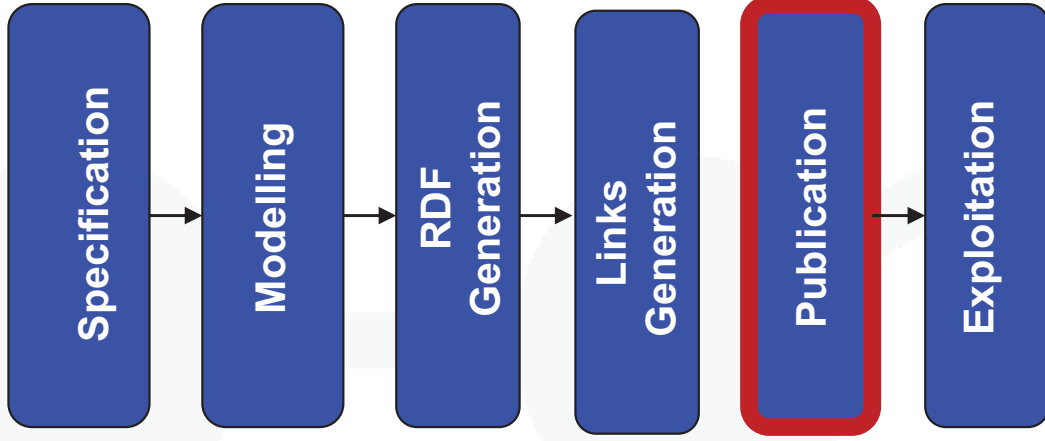
```
graph LR; A[Specification] --> B[Modelling]; B --> C[RDF Generation]; C --> D[Links Generation]; D --> E[Publication]; E --> F[Exploitation]; style D stroke:#f00,stroke-width:4px
```

Specification → Modelling → RDF Generation → **Links Generation** → Publication → Exploitation



OSW-12: Summer School on Ontological Engineering and the Semantic Web. Cercedilla. Spain





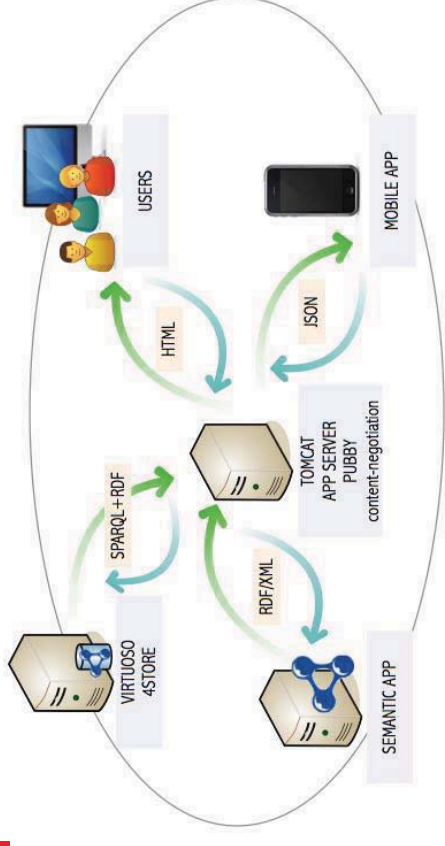
# Publicación

## Data publication

## Metadata publication using VOID

## To facilitate the discovery

- Register in CKAN your dataset
- Use to **sitemap4rdf** to generate the site map
- Upload the site map to **Google** and **Sindice**



Specification

Modelling

RDF  
Generation

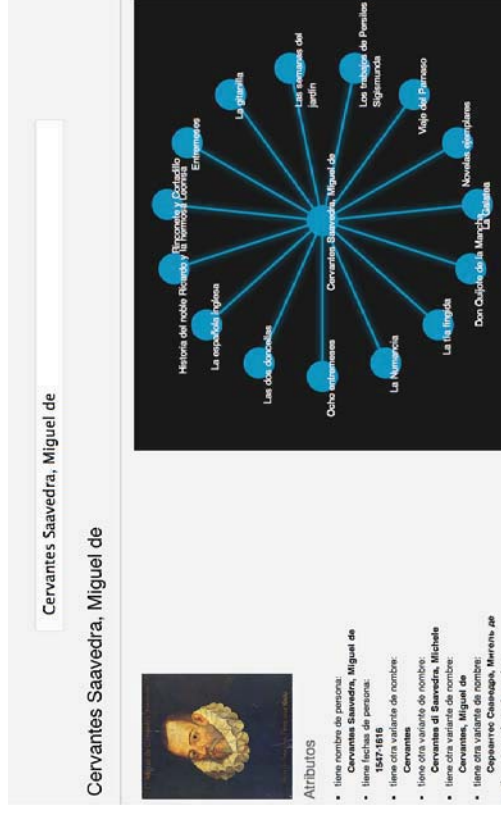
Links  
Generation

Publication

Exploitation

Exploitation

## Web Interface



## SPARQL queries

```
select distinct COUNT(?Obras) where {  
  http://datos.bne.es/resource/XX1718747  
  <http://iflastandards.info/ns/fr/frbr/frbrer/P2010>  
  ?Obras
```

URI Cervantes

Is author

<http://bne.linkeddata.es/>

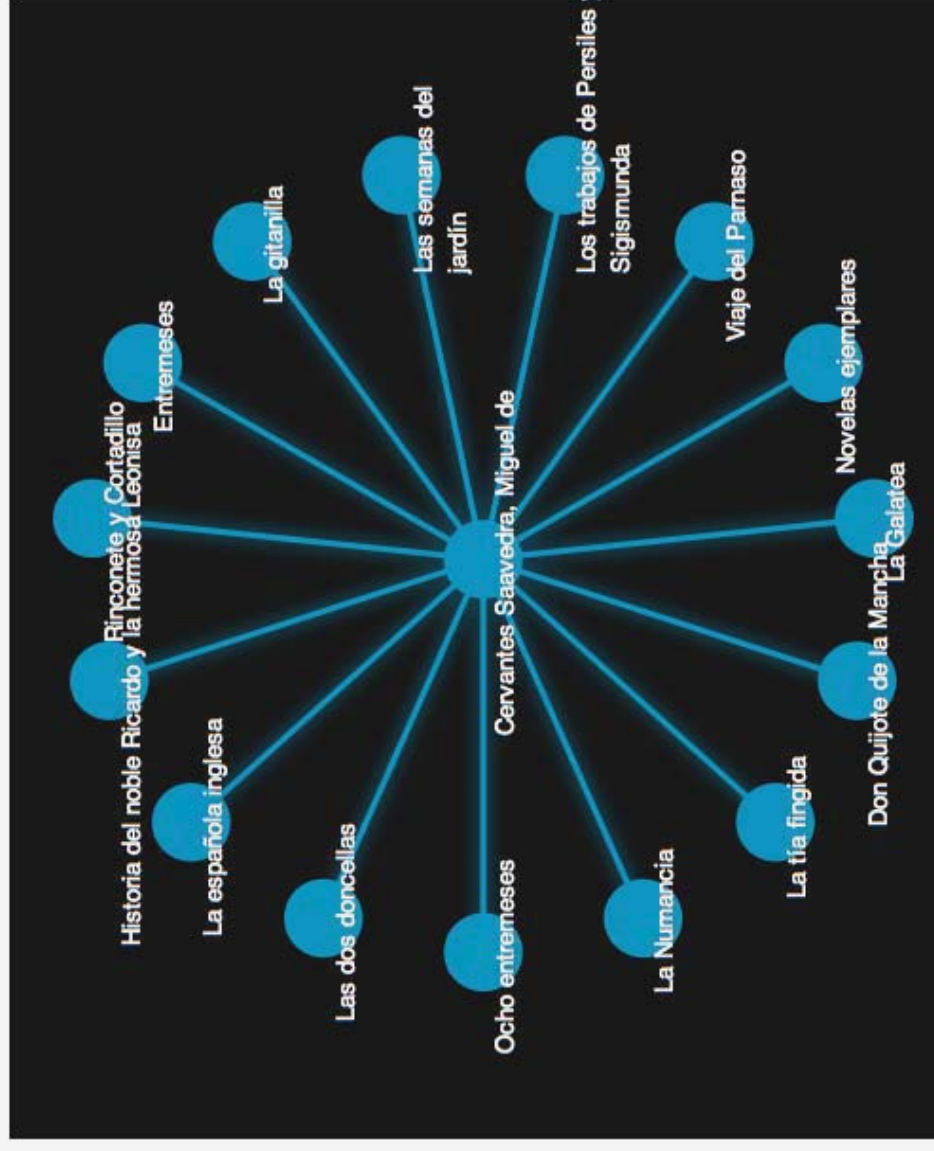
## Cervantes Saavedra, Miguel de

### Cervantes Saavedra, Miguel de



#### Atributos

- tiene nombre de persona:  
**Cervantes Saavedra, Miguel de**
- tiene fechas de persona:  
**1547-1616**
- tiene otra variante de nombre:  
**Cervantes**
- tiene otra variante de nombre:  
**Cervantes di Saavedra, Michele**
- tiene otra variante de nombre:  
**Cervantes, Miguel de**
- tiene otra variante de nombre:  
**Сервантес Сааведра, Мигель де**



Búsqueda

Seleccione campo





Todos los nombres


Seleccione índice


Todo VIAF

Términos de búsqueda

Búsqueda

Cervantes Saavedra, Miguel de, 1547-1616. | Don Quixote    


Cervantes Saavedra, Miguel de 1547-1616 | Don Quijote de la Mancha 

Cervantes Saavedra, Miguel de 1547-1616 | Don Quijote 


VIAF ID: 184295284 (Work)

Permalink: <http://viaf.org/viaf/184295284>


Formas preferidas




100 1 \_ 1a Cervantes Saavedra, Miguel de, 1d 1547-1616. 1t Don Quixote




100 1 \_ 1a Cervantes Saavedra, Miguel de, 1d 1547-1616. 1t Don Quixote




100 1 \_ 1a Cervantes Saavedra, Miguel de, 1d 1547-1616. 1t Don Quixote




100 2 \_ 1a Cervantes Saavedra, Miguel de, 1d 1547-1616 1t Don Quixote




100 1 0 1a Cervantes Saavedra, Miguel de 1d 1547-1616 1t Don Quijote de la Mancha




100 1 \_ 1a Cervantes Saavedra, Miguel de 1d 1547-1616 1t Don Quijote



100 | \_ 1a Cervantes Saavedra, Miguel de (1547-1616). 1t Don Quijote de la Mancha



• <http://www.idref.fr/027300935/id>



SSSW-12: 9th Summer School on Ontological Engineering and the Semantic Web. Cercedilla. Spain

49



bima0000029232

at datos.bne.es

http://datos.bne.es/resource/bima0000029232



Property	Value
isbd:P1004	<ul style="list-style-type: none"><li>The much esteemed history of the ever famous knight don Quixote da la Mancha</li></ul>
isbd:P1006	<ul style="list-style-type: none"><li>: containing his many wonderful adventures and atchievements, very pleasant and diverting : with the comicl humours of Sancho Pancho [sic], his remarkable'squire ...</li></ul>
isbd:P1016	<ul style="list-style-type: none"><li>[S.I.]</li></ul>
isbd:P1018	<ul style="list-style-type: none"><li>1789</li></ul>
isbd:P1022	<ul style="list-style-type: none"><li>108 p.</li></ul>
isbd:P1023	<ul style="list-style-type: none"><li>il.</li></ul>
isbd:P1024	<ul style="list-style-type: none"><li>8º</li></ul>
isbd:P1040	<ul style="list-style-type: none"><li>Port. con doble filete y grab. xil.</li></ul>
isbd:P1045	<ul style="list-style-type: none"><li>La h. de lám. es grab. xil.</li></ul>
isbd:P1064	<ul style="list-style-type: none"><li>Río y Rico</li></ul>
isbd:P1117	<ul style="list-style-type: none"><li>sold by the booksellers in Great-Britain, Ireland and America</li></ul>
isbd:P1185	<ul style="list-style-type: none"><li>[Texto impreso]</li></ul>
ifla-frbr:P2004	<ul style="list-style-type: none"><li>bne:resource/XX2261935</li></ul>
dcterms:language	<ul style="list-style-type: none"><li>&lt;http://lexvo.org/id/iso639-3/eng&gt;</li></ul>
rdf:type	<ul style="list-style-type: none"><li>ifla-frbr:C1003</li></ul>

Metadata

Anon_0	
rdf:type	<http://www.w3.org/2004/03/trix/rdfg-1/Graph>
foaf:primaryTopic	bne:resource/bima0000029232
void:subset	Anon_1 (more)
prv:createdBy	Anon_2 (more)

- **Modelling:**

- Open Metadata Registry
- Neon Toolkit



- Mapping and generation

- **MARiMbA:** Library-oriented, supports and facilitates the entire process of transformation from MARC21 to RDF



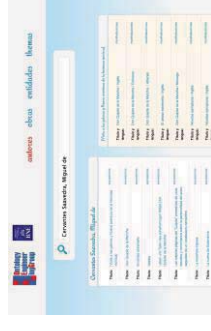
- **Publication:**

- Virtuoso Universal Server
- Pubby
- CKAN registry
- Sitemap4rdf



- **Exploitation:**

- Web Applications that visualize data using SPARQL



- Total number of authority records: **4.100.000**
- Total number of bibliographical records: **2.390.140**
- Total number of RDF triples: **58.053.215**
- Number of links: (15% authorities): **587.520**
- Linked sources:
  - VIAF
  - SUDOC (French collective university catalogue) FR
  - GND (German National Library of authorities) GER
  - LIBRIS Sweden
  - DBPedia
  - Soon BNF

<http://bne.linkeddata.es/>



## 1. The concept

## 2. Foundations

## 3. The process

## 4. Examples

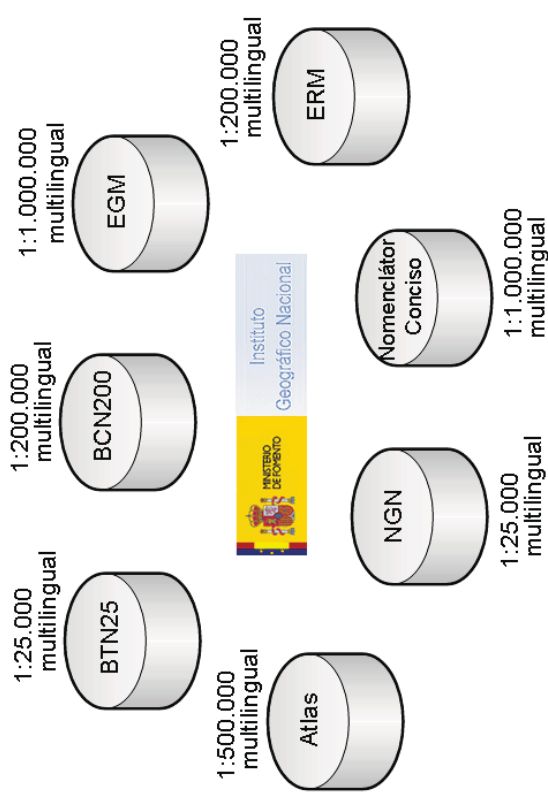
- Libraries: <http://datos.bne.es>
- <http://linkeddata3.dia.fi.upm.es/bne-demo>
- Geo: <http://geo.linkeddata.es/>
- Meteorology: <http://aemet.linkeddata.es/>
- Travelling: <http://webenemasuno.linkeddata.es/>



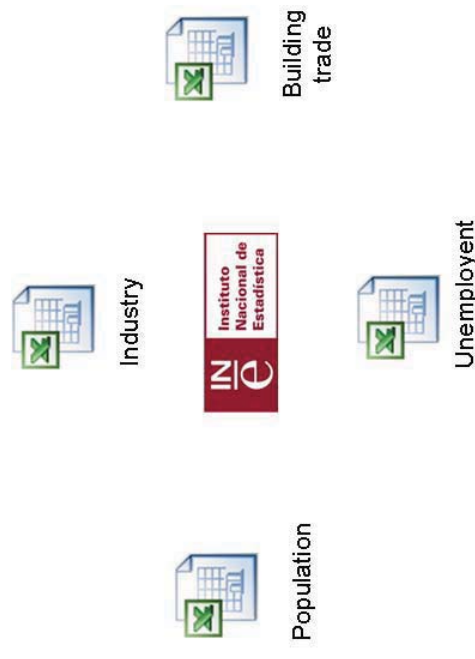


# Identification and selection of data sources

## Geographical Spanish Institute



## Statistical Spanish institute



Specification

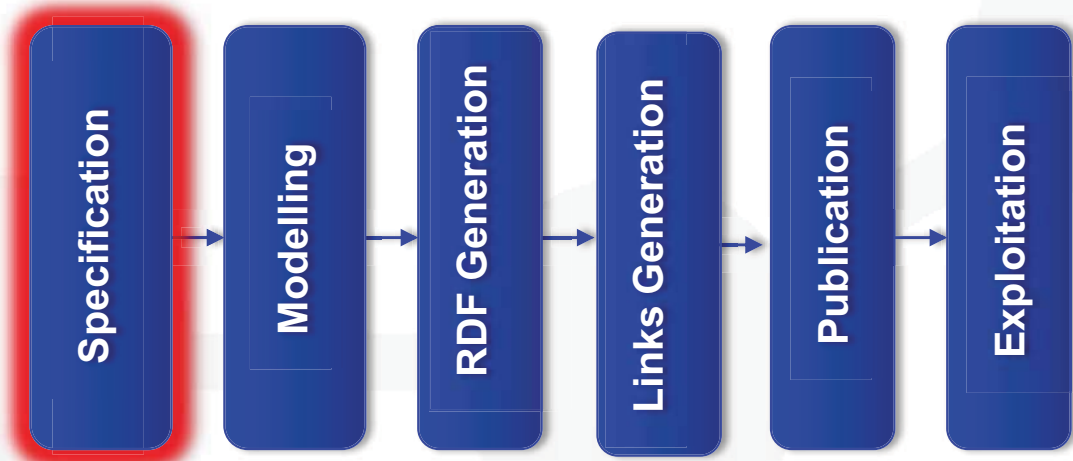
Modelling

RDF Generation

Links Generation

Publication

Exploitation



- Geographic Spanish Institute
  - Multilingual (Spanish, Vasc, Gallician, Catalan)
  - Conceptualization mismatches
  - Granularity (scale concept)
  - Textual information
    - Inform. Hidrográfica:** reservoir, river, , etc.
    - Transportes.** Vía desdoblada, Ferrocarril, ...
    - Unidades Administrativas.** Municipality
- Particularities
  - Longitude and latitude
- Statistic Spanish Institute
  - Monolingual
  - Numerical information
  - Particularities
    - Geo (textual level) and Temporal


# 1. Identification and selection of the data sources

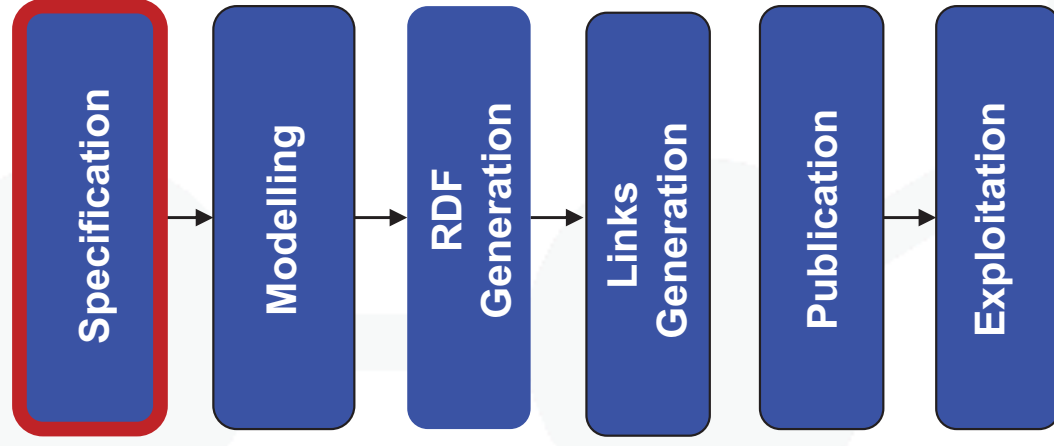
NOMBRE	ENTIDAD	COM	PROV	CODINE	XUTM	YUTM	HUSO	HOJA25	LongitudG	LatitudG	LongitudGMS	LatitudGMS
Abejuela, Olivar de / Oli	Lugar/Paraje	0	02	02042	584300	246500	30	0867-3	-2,03512445814873	38,3620693572292	-02°02'06.44	38°21'43
Abejuela, Rambla de	Corriente fluvial	0	04	04053	597560	152000	30	0974-4	-1,89618811907142	37,509177056304	-01°53'46.27	37°30'33
Abejuela, Rambla de la	Corriente fluvial	1	46	46262	579190	423000	30	0638-2	-0,90285531710738	39,9374839997029	-00°54'10.27	39°56'14
Abejuela, Rambla de la	Corriente fluvial	1	44	44002	578800	424000	30	0638-2	-0,907142388290964	39,9465700679236	-00°54'25.71	39°56'47
Abelá, A	Población	0	27	27025	521500	336500	29	0003-3	-7,49293473672413	43,670782491816	-07°29'34.56	43°40'14
Abeladoira	Población	0	27	27039	522000	774030	29	0073-1	-7,50066276897865	43,1084031553505	-07°30'02.38	43°06'30
Abeladoira	Lugar/Paraje	0	27	27039	521550	774200	29	0073-1	-7,50615404547912	43,1100057447368	-07°30'22.15	43°06'36
Abelaedo	Lugar/Paraje	0	27	27005	548100	321300	29	0010-3	-7,16730269277051	43,5291689758955	-07°10'02.28	43°31'45
Abelaedo, Chao do	Llanura/Raso	0	27	27013	525500	330200	29	0009-1	-7,4448008623332	43,613414717808	-07°26'41.28	43°36'48
Abelaedo, Monte do	Lugar/Paraje	0	27	27064	509000	341000	29	0002-4	-7,6470271512085	43,7132267731492	-07°38'49.29	43°42'47
Abelaedo, O	Población	0	27	27064	508500	340500	29	0002-4	-7,65333254552948	43,7087994633333	-07°39'11.99	43°42'31
Abelaedo, Rego do	Corriente fluvial	0	27	27064	509000	340540	29	0002-4	-7,64712030731384	43,7090862344245	-07°38'49.63	43°42'32

NOMBRE	GMLGEOMETRY
Abengibre	(HUGECLLOB)
Alatoz	(HUGECLLOB)
Albatana	
Balsa de Ves	
Ballesterio, El	
Alborea	
Alcadozo	
Alcalá del Júcar	



 <gml:Polygon srsName="SDO:8223" xmlns:gml="http://www.opengis.net/gml">

 <gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="." ts=" ">-1.48374108,39.23127677,0 -1.48404  
-1.48480765,39.22531695,0 -1.48544493,39.2231538,0 -1.4858725,39.22212319,0 -1.48608465,39.22141415,0 -1.4860584,  
-1.48421128,39.21319056,0 -1.48390698,39.21064575,0 -1.48429166,39.2094524,0 -1.48529268,39.20490629,0 -1.492506,  
-1.49590016,39.20383639,0 -1.49905691,39.20405732,0 -1.50256876,39.20446292,0 -1.50338941,39.20453654,0 -1.50483,  
-1.51378498,39.2047147,0 -1.51398161,39.20472622,0 -1.5171434,39.20471246,0 -1.51894734,39.20373528,0 -1.5196271,  
-1.52344981,39.20108038,0 -1.52699007,39.19901681,0 -1.5284903,39.19815278,0 -1.53151867,39.19619073,0 -1.532086,  
-1.5329951,39.19479473,0 -1.53398595,39.19503246,0 -1.53537021,39.19528414,0 -1.53607317,39.19544615,0 -1.536914,  
-1.53953498,39.19602121,0 -1.54099994,39.19629185,0 -1.54110394,39.19630216,0 -1.54231647,39.19646151,0 -1.54461,  
-1.54957924,39.19633592,0 -1.55131402,39.1964386,0 -1.55382202,39.19666799,0 -1.55798564,39.19690882,0 -1.559800,  
-1.56633376,39.19750743,0 -1.5703604,39.19765603,0 -1.57231502,39.19777011,0 -1.57278711,39.19790204,0 -1.573636,  
-1.57633787,39.19982863,0 -1.57807687,39.20087703,0 -1.57814542,39.20092292,0 -1.57871557,39.20136237,0 -1.57915,  
-1.5805024,39.20233924,0 -1.58088715,39.20278543,0 -1.58107708,39.20313014,0 -1.58122865,39.20365457,0 -1.581447,  
-1.5819531,39.20563664,0 -1.58191475,39.20581637,0 -1.58194513,39.20603299,0 -1.58285235,39.20756676,0 -1.584270,  
-1.48544493,39.2231538,0 -1.4858725,39.22212319,0 -1.48608465,39.22141415,0 -1.4860584,39.22531695,0 -1.48374108,39.23127677,0 -1.48404

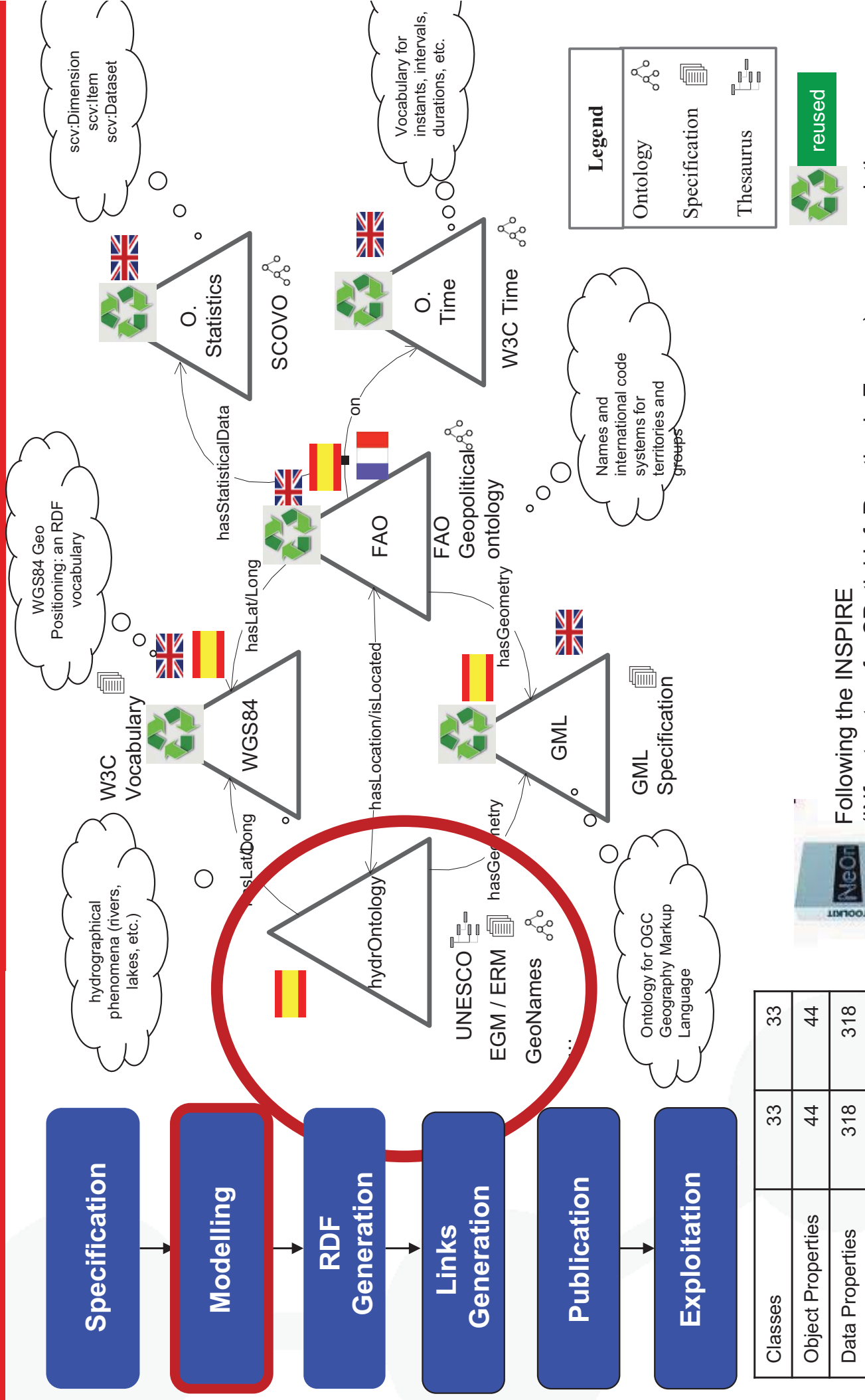


POBLA\_01  
 POBLA\_08  
 PBL01\_08  
 TS\_CREC\_EF  
 DENS\_01  
 DENS\_08  
 DENS01\_08  
 TS\_CREC\_VE  
 TS\_CREC\_VE  
 TS\_CREC\_VE  
 TS\_NAT\_01  
 TS\_NAT\_08  
 TS\_NAT01\_0  
 TS\_MORT\_01  
 TS\_MORT\_08  
 TS\_MORTAL\_  
 TS\_FECUN\_0  
 TS\_FECUN\_0  
 TS\_FECUN01  
 MUJ\_FERTIL  
 MUJ\_FERTIL  
 MUJ\_FETIL\_  
 POBL\_EXTR\_  
 POBL\_EXTR\_  
 EVOL\_POBL\_  
 IND\_DEP\_01  
 ÷ND\_DEP\_08  
 IND\_DEPEN\_  
 ESP\_IND\_EN  
 ESP\_÷NDI\_E  
 ESP\_IND\_EN  
 PORT\_IND\_E  
 PORT\_÷NDI\_

Población año 2001  
 Población año 2008  
 Evolución de la población desde el 2001 al 2008  
 Tasa de Crecimiento efectivo  
 Densidad de Población año 2001  
 Densidad de Población año 2008

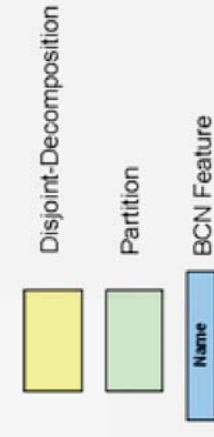
	2009	2008	2007	2006	2005	2004	2003
Total Nacional	3355830	3422239	3336657	3174393	3064129	2942593	2813159
Alava	21988	22318	20676	20349	19838	19779	19638
Albacete	27380	27647	27068	25531	24685	23550	22547
Alicante	136239	142307	140145	133016	123333	113852	111805
Almería	43501	45130	43970	40871	38766	36260	33947
Asturias	71853	73124	72276	70115	68175	67039	65062
Ávila	11455	11708	11434	10900	10611	10319	10211
Badajoz	40874	41358	40168	38045	37052	34972	34866
Illes Balears	91826	93335	91254	88027	87024	85425	75951
Barcelona	467365	477942	469432	444410	436294	417425	397693
Burgos	25567	25891	25372	24504	23733	22882	22159
Cáceres	26307	26494	26064	25039	24846	20596	23440
Cádiz	62817	64505	63338	61691	58986	57138	54462
Cantabria	39611	40393	39560	37690	36561	35649	34017
Castellón	42122	43855	42476	39749	37865	37214	34213
Ciudad Real	32046	33011	31881	30446	29521	29011	26778
Córdoba	48979	50057	49302	47155	45405	43394	41964
Coruña, A	83748	84220	82873	79170	77023	74809	71748
Cuenca	14747	14928	14741	13822	13336	12829	12546
Girona	58404	51467	50108	47169	46827	45145	52482
Granada	60016	62269	61055	57223	54341	50508	49862
Guadalajara	13507	13735	12874	11825	10438	10120	9422
Guipúzcoa	62034	63569	59546	58486	57193	56498	55983
Huelva	26783	27463	27063	25487	24777	24270	22547
Huesca	16837	17109	16694	16025	15390	15078	14283
Jaén	36557	37368	36962	35383	34675	33157	32444
León	33564	34012	33563	32359	31664	30992	30256
Lleida	36920	37638	36065	33956	32739	31515	29605
Lugo	24861	25035	24609	23780	23122	22479	22396
Madrid	511804	519307	503000	478202	456175	436074	407655
Málaga	113362	116683	114547	108713	102382	96587	88257
Murcia	95636	100075	97374	90698	85110	82484	75973
Navarra	43282	43847	43142	41083	40730	39679	38936
Ourense	23304	23711	23520	22843	22452	22118	21560
Palencia	10664	11144	11060	10694	10575	10399	10297

## 2. Lightweight Ontology Development

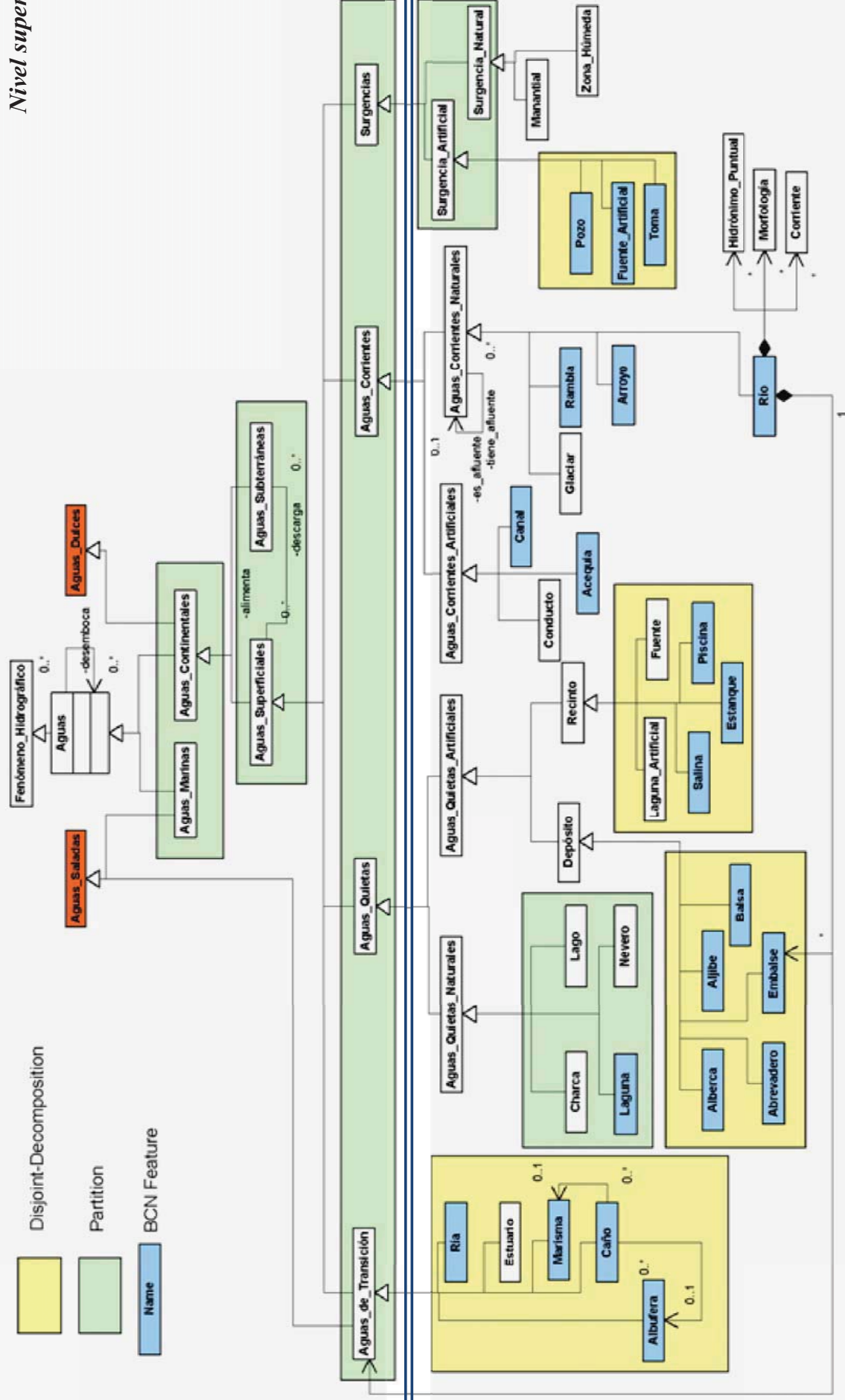


Following the INSPIRE (Infrastructure for Spatial Information in Europe) recommendation. hydroOntology, SCOVO, FAO Geopolitical, WGS84, GML, and Time





Nivel superior



Nivel inferior